Am. J. Hum. Genet. 57:201–223, 1995

1994 WILLIAM ALLAN AWARD ADDRESS
Mitochondrial DNA Variation in Human Evolution, Degenerative Disease, and Aging

Douglas C. Wallace

Department of Genetics and Molecular Medicine, Emory University School of Medicine, Atlanta

The delineation of human mtDNA variation and genetics over the past 25 years has provided unique and often startling new insights into human evolution, degenerative diseases, and aging. While the autosomal nuclear DNA (nDNA) genes are confined to the nucleus, limited to two copies per cell, and transmitted equally from both parents according to the laws of Mendel, the mtDNA genes are disbursed throughout the cytoplasm, present in hundreds to thousands of copies per cell, and transmitted exclusively through the mother. While the biochemical defects and clinical phenotypes of nuclear mutations are quantized and frequently stereotyped, the proportion of mutant and normal mtDNAs fluctuates between daughter cells, siblings, and even identical twins, resulting in continuous biochemical defects and variable phenotypes.

Thus, while the strict maternal inheritance of the mtDNA means that the mitochondrial genome evolves by simple sequential mutations, the stochastic distribution of mutant mtDNAs during cell division adds considerable complexity to the familial transmission and expression of genetic disease. Consequently, the mtDNA provides a simple system for reconstructing ancient human migrations but also contributes to the variability of human genetic diseases, partially explaining variable expressivity and penetrance, delayed onset of symptoms, and even variable rates of aging.

Human Mitochondrial Genetics

The mtDNA is a 16,569–nucleotide pair (np) closed circular molecule located within the matrix of the double-membrane mitochondrion. Each human cell contains hundreds of mitochondria and thousands of mtDNAs, and each mtDNA codes for 13 polypeptides essential to the enzymes of the mitochondrial energy-generating pathway oxidative phosphorylation (OXPHOS), plus the small (12S) and large (16S) rRNAs and 22 tRNAs necessary for mitochondrial protein synthesis (Anderson et al. 1981; Wallace et al. 1995). OXPHOS is located within the mitochondrial inner membrane and is composed of five multiple polypeptide enzyme complexes whose genes are dispersed between the mtDNA and the nDNA. The mtDNA encodes seven (ND1, -2, -3, -4, -4L, -5, and -6) of the ~42 polypeptides of OXPHOS complex I (NADH: ubiquinone oxidoreductase), 1 (cytochrome b, cytb) of the 11 polypeptides of complex III (ubiquinol: cytochrome c oxidoreductase), 3 (COI, COII, and COIII) of the 13 polypeptides of complex IV (cytochrome c oxidase), and 2 (ATP6 and 8) of the 16 polypeptides of complex V (proton-translocating ATP synthase) (fig. 1).

OXPHOS generates ATP by oxidizing hydrogens derived from the metabolism of carbohydrates and fats with O_2 to give H_2O. The electrons from the hydrogens are transferred to complex I of the electron transport chain via NADH. From complex I, the electrons are transferred to ubiquinone (CoQ_{10}), then complex III, cytochrome c, complex IV, and finally to O_2. The energy that is released is used to pump protons across the mitochondrial inner membrane through complexes I, III, and IV, creating an electrochemical gradient ($\Delta\Psi$) with excess protons (H^+) on the outside and hydroxyl ions (OH^-) on the inside. The $\Delta\Psi$ is utilized by complex V as a source of potential energy to condense ADP + Pi to make ATP, and the resulting matrix ATP is exchanged for spent cytosolic ADP by the inner membrane adenine nucleotide translocator (Wallace 1992).

The endosymbiotic origin (Wallace 1982*b*), cytoplasmic location, and high copy number of the mitochondria make for a unique genetics. Even though the mitochondria have become integral components of the human cell, the mitochondria retain a semiautonomous genetic system. This was first demonstrated in the early 1970s, using cultured mammalian cells resistant to the mitochondrial ribosome inhibitor chloramphenicol (CAP), by showing that CAP resistance could be transferred between cells by using our cytoplasmic hybrid (cybrid) system (fig. 2) (Bunn et al. 1974; Wallace et al.

Received May 19, 1995. This article was adapted by the author from the text of his 1994 William Allan Award Address, delivered October 22, 1994, in Montreal. It represents the opinions of the author and has not been peer reviewed.

Address for correspondence and reprints: Dr. Douglas C. Wallace, Department of Genetics and Molecular Medicine, Emory University School of Medicine, 1462 Clifton Road, N.E., Atlanta, GA 30322.

0002-9297/95/5702-0001$02.00

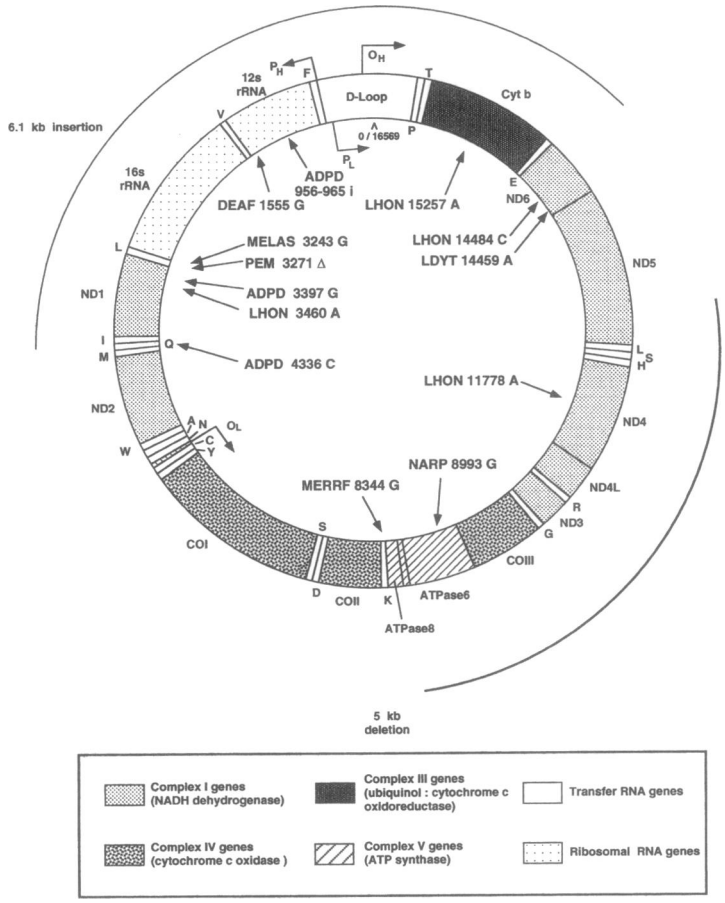

Figure 1 Human mtDNA structural and morbid maps. The polypeptide and rRNA gene symbols are defined in the text. The tRNAs are indicated by their cognate amino acid single letter code. The genes encoded by the G-rich heavy (H) strand are on the outside of the circle, while those for the C-rich light (L) strand are on the inside. The H- and L-strand origins (O_H and O_L) and promoters (P_H and P_L) are shown. The positions of representative base substitutions are shown on the inside of the circle. The disease acronyms are defined in the text, with the exception that ADPD stands for Alzheimer and Parkinson disease and DEAF for neurosensory hearing loss. Two representative rearrangements are shown outside the circle, the common 5-kb deletion associated with ocular myopathy and Pearson marrow/pancreas syndrome and the 6.1-kb insertion observed in the maternally inherited diabetes and deafness family harboring the trimolecular heteroplasmy (Anderson et al. 1981; Wallace et al. 1995).

1975). CAP-resistant (CAP-R) cells were enucleated by treatment with cytochalasin B, which disaggregates the cytoskeleton, plus suspension in a centrifuged field. This pulls the dense nucleus out of the cell until it is connected to the cytoplasm by a thin cytoplasmic filament. When this filament breaks, independent membrane-bound nuclei and cytoplasms (cytoplasts) are generated. The cytoplasts contain the mitochondria and mtDNAs, and when cytoplasts from CAP-R cells are fused to a CAP-sensitive (CAP-S) recipient cells, the resulting cybrids acquire the CAP resistance of the cytoplasmic donor (Bunn et al. 1974; Wallace et al. 1975). Subsequent analysis showed that CAP-resistance was cotransferred with mtDNA RFLPs (Wallace 1981) and mitochondrial protein polymorphisms (Oliver and Wallace 1982), and mtDNA sequence analysis revealed that CAP resistance was the product of single nucleotide changes in the 16S rRNA gene (Blanc et al. 1981*a*, 1981*b*; Kearsey and Craig 1981).

The human mtDNA is additionally unique in being maternally inherited. Using RFLPs as mtDNA genetic markers in family studies, we showed that the mtDNA was transmitted from the mother to all of her offspring and from her daughters to the next generation, but males did not transmit their mtDNA (Giles et al. 1980; Case and Wallace 1981).

Another novel feature of mitochondrial genetics is that mitochondria behave as an intracellular population of microorganisms. The population dynamics of the intracellular mitochondria was first explored using CAP-R and CAP-S mammalian cells (Bunn et al. 1977; Wallace et al. 1977; Wallace 1982*a*, 1986). When a mutation arises in one of the mtDNAs within a cell, this creates an intracellular mixture of mutant and normal

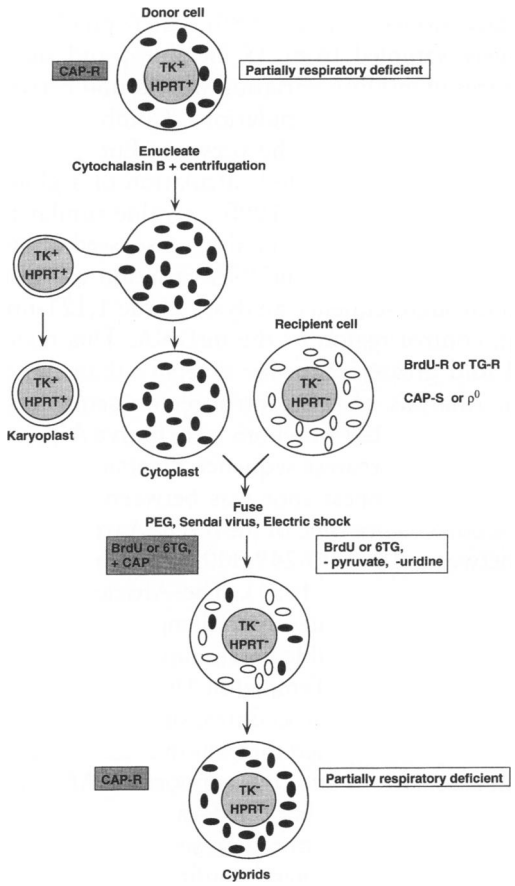

Figure 2 Cybrid-transfer procedure for identifying mtDNA mutations. Cybrid transfer has been used for demonstrating the cytoplasmic transfer of CAP resistance and for linking clinically important mtDNA mutations to mitochondrial OXPHOS defects. This drawing shows the successive stages of cybrid fusion. The mtDNA donor cell is enucleated and fused to the recipient cell. The cybrids are isolated by selecting for the recipient cell nucleus and for the donor cell mtDNAs. For transfer of CAP resistance (shaded boxes on left of cells), the CAP-R cytoplasts are fused to recipient cells containing CAP-S mtDNAs, and the cybrids are selected in BrdU or 6TG plus CAP. For transfer of disease-associated mtDNA mutations, the patient cytoplasts are fused to cells that have been cured of their mtDNAs (ρ^0). The cybrids are selected in BrdU or 6TG plus media deficient in uridine and/or pyruvate. Since ρ^0 cells die without these two metabolites, only the cybrids with partially functional respiratory chains will grow.

molecules, known as heteroplasmy. When a heteroplasmic cell divides, it is a matter of chance which mtDNAs will be partitioned into the daughter cells. As a result, over many cell divisions the proportion of mutant and normal mtDNAs can drift toward either pure mutant or pure normal (homoplasmy), a process known as replicative segregation. Since replicative segregation can occur during the replication of somatic cells or during the proliferation of female germinal cells leading to oocytes, it can result in differences in the proportion of mutant mtDNAs within the tissues of a heteroplasmic individual or among the offspring of a heteroplasmic mother. For deleterious mtDNA mutations, as the proportion of mutant mtDNAs increases, the mitochondrial ATP-generating capacity declines. Ultimately, the energy output becomes insufficient to sustain cell and tissue function (the bioenergetic expression threshold) and symptoms ensue. Different tissues have different bioenergetic thresholds, with the brain being most reliant on mitochondrial energy, followed by heart and skeletal muscle, kidney, and endocrine systems. Therefore, as the percentage of mutant mtDNAs fluctuates, the nature and severity of clinical symptoms varies (Wallace 1992*a*, 1992*b*).

Finally, the mtDNA is unique in that it has a very high sequence evolution rate (mutation times mutation fixation), perhaps 10–20 times higher than that of nuclear genes of comparable function (Neckelmann et al. 1987; Wallace et al. 1987). Thus, gene-for-gene, it is much more likely to acquire a deleterious mtDNA mutation than nDNA mutation. The significance of each mtDNA mutation to human variability and disease depends on where in the genome and when in the human life cycle the mutation occurs. Germ-line mtDNA mutations can be either neutral or deleterious. Neutral and mildly deleterious mutations have accumulated along human female lineages for tens of thousands of years, and, since the mtDNA is uniparentally inherited, mtDNAs rarely if ever mix and recombine. Thus, the number of mtDNA sequence differences that separate two individuals is directly proportional to the time since they shared a common maternal ancestor. Moderate to severely deleterious mutations, by contrast, cause disease and thus are eliminated by natural selection. Hence, moderately deleterious mtDNA mutations are of necessity recent events and thus are frequently heteroplasmic. Finally, somatic mutations accumulate in the mtDNAs of postmitotic tissues as we age. As they occur, they appear to erode organ and tissue function and thus may be important factors in the organ senescence associated with aging (Wallace 1994).

mtDNA Variation and Human Origins

The nature, prevalence, and geographic distribution of neutral and near neutral mtDNA variation reflects the prehistory of women. Since mtDNA mutations accumulate sequentially along radiating female lineages, individual mtDNA lineages have diverged as women migrated from Africa and into the various continents during the ~150,000 years of human evolution.

The world mtDNA phylogeny and the origin of women.— The first clear evidence that mtDNA variation correlated with the ethnic and geographic origin of the individual came from our survey of *Hpa*I RFLPs in African, Asian, and European-American mtDNAs. This revealed that in Africans, 96% of Pygmies, 93% of San Bushman, and 71% of Bantus harbored an *Hpa*I restriction site at np

3592 not seen in Asians or Europeans. The *Hpa*I np-3592 site is defined by the 5′ end of the recognition sequence, but it is caused by a C-to-T transition at np 3594. By contrast, ~13% of Asians lacked an *Hpa*I restriction site at np 12406 (G to A at np 12406), which was present in all other mtDNAs (Denaro et al. 1981). A further survey of the mtDNA variation, detected using six highly informative restriction enzymes (*Hpa*I, *Bam*HI, *Hae*II, *Msp*I, and *Ava*II) and Southern blotting, confirmed that mtDNA variation was high and correlated strongly with the geographic origin of the individual. It also showed that all mtDNAs were part of a single phylogenetic tree, that the greatest variation was in Africa, and that the tree was ~100,000 years old (Johnson et al. 1983). Extensive studies by our group, as well as by others, ultimately led to the characterization of 3,065 mtDNAs from 62 geographic samples using these six enzymes. This revealed 149 haplotypes and 81 polymorphic sites. This analysis confirmed that the mtDNA polymorphisms within each mtDNA were in virtually total linkage disequilibrium, consistent with a low frequency of recombination; that mtDNA variation correlated highly with the ethnic and geographic origin of the individual; that there was a single mtDNA tree; and that the greatest variation and deepest root of the tree was in Africa, consistent with an African origin of humans. The extent of mtDNA sequence differences between continental populations was estimated from this data by calculating the G_{ST} statistic. For the mtDNA, the G_{ST} was 0.35 ± 0.025, implying that ~35% of the mtDNA variation was continent specific. By contrast, the comparable nDNA value was 0.12. Hence, the mtDNA encompasses much greater continent-specific sequence diversity than the nDNA (Merriwether et al. 1991).

A recent African origin of human mtDNAs has also been demonstrated by the concurrent investigations of Wilson, Cann, Stoneking, and coworkers (for discussion, see Cann et al. 1987). These investigators purified the individual mtDNAs from cells or tissues, digested the DNA with 12 restriction endonucleases (*Hpa*I, *Ava*II, *Fnu*DII, *Hha*I, *Hpa*II, *Mbo*I, *Taq*I, *Rsa*I, *Hin*fI, *Hae*III, *Alu*I, and *Dde*I), end-labeled the fragments, and resolved the fragments using polyacrylamide gels and autoradiography (Brown 1980). A survey of 147 mtDNAs, including 34 Asians, 21 Australian aboriginals, 26 aboriginal New Guineans, 46 Caucasians, and 20 Africans (18 of whom were Black Americans), revealed that there was a single mtDNA tree, that the deepest root occurred in Africa, and that Africa harbored the greatest sequence diversity. Hence, Africa is the origin of *Homo sapiens*. Using an estimated sequence evolution rate of 2%–4%/million years (MYR), the human mtDNA tree was calculated to be ~200,000 years old (Cann et al. 1987).

This analysis was extended to include 62 Japanese (Horai and Matsunaga 1986) and 119 Papua New Guineans (Stoneking et al. 1990). The Papua New Guineans were sampled from 25 localities, and significant differences in mtDNA variation were found between the highland and coastal populations. Combing the Papua New Guinea data with the previous European, Asian, and African data permitted calculation of a global G_{ST} of 0.31 (Stoneking et al. 1990), a value similar to that found for the six-enzyme analysis discussed above.

The African origin of mtDNA variation was also supported through sequence analyses of the 1,121-np, noncoding, control region of the mtDNA. This region has a 3–4-fold greater sequence diversity than the coding region. Analysis of the control-region sequences from 189 individuals, 121 of whom were native African, confirmed that the greatest sequence diversity was in Africans, that the deepest root was between Africans, and that the coalescence time of the mtDNA tree (phylogeny) was between 166,000-249,000 years before present (YBP) (Vigilant et al. 1991). The African root of this phylogeny was subsequently challenged on the basis that multiple equally probable parsimony trees could be generated from the data (Templeton 1992). However, other phylogenetic analysis procedures, such as neighbor-joining trees, have reaffirmed the cohesiveness of the deepest African associations and thus support the African origin of the mtDNA phylogeny (Hedges et al. 1991).

Analysis of the control-region sequence of 95 individuals, including 61 Japanese, confirmed that the greatest diversity and deepest root occurred in Africa and revealed that "Mongoloid" mtDNAs were subdivided into two distinct groups (Horai and Hayasaka 1990). Analysis of 117 Caucasian mtDNAs confirmed the distinctive nature of many European mtDNAs and revealed that the various mtDNA lineages were widely disseminated throughout Europe (DiRienzo and Wilson 1991).

Finally, comparison of the original European mtDNA sequence (Anderson et al. 1981) with that from an African, a Japanese, and four African apes (common and pygmy chimpanzees, gorilla, and orangutan) revealed that the European and Japanese mtDNAs were most similar, that the African mtDNA was more divergent, and that the nearest ape relatives, the chimpanzees, were 10 times more divergent from humans than Africans are from Asians and Europeans. Using the orangutan-African ape divergence time of 13 million years before present (YBP) as reference, this study gave an age for human mtDNA radiation of 143,000 ± 18,000 YBP and a time for European and Japanese radiation of 70,000 ± 13,000 YBP (Horai et al. 1995).

All of these studies lead to the same conclusion. The human mtDNA tree appears to have originated in Africa ~150,000 YBP. As women migrated from Africa to colonize new lands, additional mtDNA mutations arose and became established by genetic drift, resulting in continent-specific mtDNA variation. Today, these popula-

tion-specific polymorphisms constitute the background on which potentially pathogenic mtDNA mutations must be identified.

Cataloging continent-specific mtDNA variation.—While the above methods permitted elucidation of the general features of human mtDNA evolution, a more detailed analysis of mtDNA variation was necessary for clinical studies and for addressing additional anthropological questions on the age and origin of Africans, Europeans, Asians, and Native Americans. To increase the sensitivity of our analyses, we developed a new mtDNA analysis procedure in which the mtDNAs from a variety of human samples could be amplified by using PCR in nine overlapping fragments. Each fragment was then digested with 14 restriction endonucleases (*Alu*I, *Ava*II, *Bam*HI, *Dde*I, *Hae*II, *Hha*I, *Hin*fI, *Hinc*II, *Hpa*I, *Hpa*II/*Msp*I, *Mbo*I, *Rsa*I, and *Taq*I), and the fragments resolved on agarose gels and detected by ethidium bromide staining and UV fluorescence. This procedure scans >20% of the mtDNA sequence, and the aggregate of the restriction-site polymorphisms for each mtDNA defines the mtDNA haplotype (Ballinger et al. 1992*a;* Torroni et al. 1992). The regional PCR fragments can also be sequenced, permitting extension of the analysis to areas of interest such as the hypervariable control region (Torroni et al. 1993*a,* 1993*b*).

The sequence differences between mtDNAs could then be compared by using various phylogenetic procedures including parsimony, neighbor-joining, and unweighted pair-group analyses. These phylogenetic trees reveal the relatedness of the mtDNAs, with the more similar mtDNAs clustering together. The extent of sequence diversity within or between groups of related haplotypes (haplogroup) can also be calculated (Tateno et al. 1982; Nei and Tajima 1983; Saitou and Nei 1987; Swofford 1993).

African mtDNA variation.—To better characterize African mtDNA variation, we have surveyed the mtDNAs from 140 Africans, including 101 from Senegal, of whom 60 were Mandenkalu, 20 Wolof, 8 Pular, and 13 others from eight tribes. In addition, we analyzed 22 Eastern Pygmies from Zaire and 17 Western Pygmies from Central African Republic (Chen et al. 1995). This survey revealed 79 haplotypes defined by 119 polymorphic sites. Phylogenetic analysis revealed that 55 of the haplotypes formed a single, coherent, African-specific haplogroup designated "L" (fig. 3), which is defined by the African-specific *Hpa*I site at np 3592 together with the *Dde*I site at np 10394 (A to G at np 10398). This lineage is subdivided into two sublineages, L1 and L2. L1 encompasses 39% of the L haplotypes and 34% of all African mtDNAs and is defined by an additional *Hin*fI site at np 10806 (T to C at np 10810). L2 encompasses 61% of the L haplotypes and 42% of the African mtDNAs and is defined by an additional combined *Hin*fI site gain at np 16389 and *Ava*II site loss at np 16390 (G to A at np 16390). All L haplotypes are of African origin, with the only exceptions occurring in populations known from historical evidence to have had African contact.

Several other features of haplogroup L are of interest. Two length mutations have been observed in L1: a 9-np COII/tRNA^{Lys} deletion between nps 8272 and 8289 (Cann and Wilson 1983; Wrischnik et al. 1987) found in two African haplotypes, AFR 60 (representing 27% of Eastern Pygmies) and AFR 61 (representing 24% of Western Pygmies); and a 10–12-bp insertion of cytosines (Cs) between the tRNA^{Tyr} and COI gene (nps 5895–5899) in three Western Pygmies with the AFR66 haplotype (fig. 3).

While the West African mtDNAs are dispersed throughout subgroups L1 and L2, the Eastern and Western Pygmies are asymmetrically distributed between these two subhaplogroups. Close to 65% of Western Pygmy mtDNAs clustered in a single L1 sublineage defined by a *Taq*I site at np 9070 and an *Rsa*I site at np 12810, whereas 54% of Eastern Pygmy mtDNAs clustered in L2 sublineage defined by a *Dde*I site loss at np 13065 and a *Rsa*I site at np 11776.

The remaining African mtDNAs form a heterogeneous array of five lineages, each defined by specific restriction-site gains or losses. While the L lineage is well defined and is a consistent feature of African phylogenies (see consensus tree in the insert of fig. 3) the other lineages are less consistently defined. One of these lineages is defined by loss of the *Dde*I site at np 10394. This lineage represents only 4% of the African mtDNAs, yet it appears to be the progenitor of roughly half of all European, Asian, and Native American mtDNAs. Within this lineage are mtDNAs that also lack a *Hin*fI site at np 12308. This mtDNA haplotype is closely related to the European-specific haplogroup N.

Calculation of the accumulated sequence diversity of the African-specific haplogroup L and its subhaplogroups L1 and L2 gave values of 0.285%, 0.249%, and 0.172%, respectively. The total African mtDNA sequence diversity was 0.292%. This means that haplogroup L has the highest sequence diversity of any continent-specific haplogroup and that Africa encompasses the greatest diversity of any continent. Using our estimate of the mtDNA sequence evolution rate of 2.2%–2.9%/MYR (Torroni et al. 1994*d*), the L haplogroup is between 98,000–130,000 years old, and the total African mtDNA lineage is between 101,000–131,000 years old. (fig. 4)

European mtDNA variation.—To refine our description of European mtDNA sequence variation, we collected and analyzed 175 samples from individuals of European ancestry living in the United States and Canada, here designated Europeans (Torroni et al. 1994*b*). Restriction

Figure 3 Maximum parsimony tree of African mtDNA haplotypes. The large phylogeny is one example of a maximum parsimony tree generated by the PAUP program (Swafford 1993) and rooted using an Asian mtDNA as the outgroup. The insert is a consensus tree derived from 1,000 maximum parsimony trees after 24,012 replications. The primary branches of the main tree (L, L1, and L2) are indicated, as are the haplotypes of the Eastern and Western Pygmies, the two haplotypes associated with the 9-bp deletion (AFR60 and 61), and the haplotype baring the 12-bp insertion (AFR60). The numbers at the nodes indicate important restriction site gains (+) or losses (−), with the nucleotide number representing the 5′-end of the restriction site on the L-strand. The letter after the nucleotide number is the relevant restriction enzyme: a, *Alu*I; b, *Ava*II; c, *Dde*I; e, *Hae*III; g, *Hin*fI; h, *Hpa*I; j, *Mbo*I; k, *Rsa*I; and l, *Taq*I. This figure is reprinted from (Chen et al. 1995, with permission).

analysis revealed 117 haplotypes, defined by 137 polymorphic sites. Only 19 of these haplotypes were found in more than one individual, with haplotype 4 representing 13% of European mtDNAs. Phylogenetic analysis revealed that all European mtDNAs were subdivided into two groups by the presence (1/4) or absence (3/4) of the *Dde*I site at np 10394 (fig. 5). Thus, Europeans exhibit a marked increase in the proportion of *Dde*I np 10394 minus mtDNAs over the 4% seen in Africans.

In addition to the macrosubdivision of European mtDNAs by the *Dde*I site at np 10394, four distinct European mtDNA haplogroups have also been observed: H, I, J, and K (fig. 3). Haplogroup H lacks the np 10394 *Dde*I site and an *Alu*I site at np 7025 (C to T at np 7028). This haplogroup encompasses 40% of European mtDNAs, and its nodal haplotype is 4. Haplogroups I, J, and K all have the np 10394 *Dde*I site. Haplogroup I encompasses 7.4% of European mtDNAs, and is defined by the *Dde*I np 1715 and HaeII np 4529 site losses, an *Alu*I np 10028 site gain, a combined *Ava*II

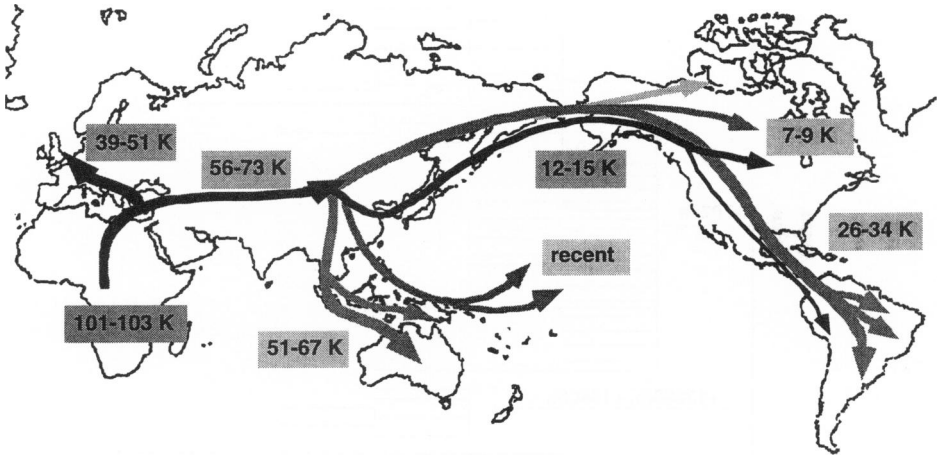

Figure 4 Sequence and timing of ancient human mtDNA radiation. The primary dates presented were calculated using 2.2%–2.9%/MYR as the evolution rate. *K* = thousands of years before present.

np 8249 site gain and *Hae*III np 8250 site less, and a combined *Bam*HI/*Mbo*I np 16389 site gain and *Ava*II np 16390 site loss. Haplotype J encompasses 9.1% of European mtDNAs and is defined by *Bst*NI np 13704 and *Hin*fI np 16065 site losses. Finally, haplotype K represents 7.4% of European mtDNAs and is delineated by a combined *Hae*II np 9052/*Hba*I np 9053 site loss. In total, haplogroups H, I, J, and K account for 64% of all European mtDNAs.

From the sequence diversity of each haplogroup their approximate ages were calculated, using the 2.2%–2.9% estimate for the mtDNA sequence evolution rate. Haplogroup H arose between 31,000–41,000 YBP, haplogroup I between 26,000–34,000 YBP, and haplogroups J and K arose between 13,000–19,000 YBP. The overall sequence divergence between the two major branches of the European phylogeny is 0.113%, giving an age for the colonization of Europe of between 39,000–51,000 YBP (Torroni et al. 1994*b*) (fig. 4).

Analysis of the control regions of the European mtDNAs revealed additional continent-specific markers (DiRienzo and Wilson 1991). However, one control-region mutation in haplogroup I proved to be totally novel, with potential implications for the evolution of the human mtDNAs. All haplogroup I mtDNAs were found to have a homoplasmic insertion of 2–6 Cs within a cluster of Cs in the sequence ACCCCCC (Box 2), where the A is located at np 567. This germ-line mutation increases the homology between this sequence and the nearby control-region sequence ACCCCCCCTCCCCCGCT (Box 1), A at np 302. Because of this homology, every individual who inherits the Box 2 germ-line insertion mutation becomes prone to undergo a somatic mutation during development. In this somatic mutation the region between Boxes 1 and 2 is duplicated as a 270-np direct repeat, possibly through slipped misreplication (Torroni et al. 1994*b*). The somatic 270-bp duplication duplicates the H-strand promoter (nps 545–567), the L-strand promoter (nps 392–445), the two intervening mitochondrial transcription factor binding sites (nps 418–445 and nps 523–550), conserved-sequence box (CSB) 3 (nps 346–363), part of CSB2 (nps 299–315), and the putative replication primer processing site (nps 317–321) (fig. 1) (Torroni et al. 1994*b*; Brockington et al. 1993). This raises the possibility that the duplicated molecules are transcribed twice as frequently and may be preferentially replicated, providing selective advantage for this mutation. This may explain why the unstable Box 2 insertion has been maintained throughout the 34,000-year history of haplogroup I (Torroni et al. 1994*b*).

Asian mtDNA variation.—To further define Asian mtDNA sequence variation, we analyzed the mtDNAs from 153 Central and Southeast Asians, including aboriginal Malays and Orang Asli, aboriginal Borneans, Han Chinese, Vietnamese, Koreans, and Malaysian Indians (Ballinger et al. 1992*a*) as well as 54 Tibetans (Torroni et al. 1994*c*) and 411 Siberians from 10 aboriginal populations, including the Chukchi and Koryaks from northeasternmost Siberia (Torroni et al. 1993*b*). The Asian phylogenetic tree encompasses 42 Tibetan haplotypes, 106 Asian haplotypes, and 34 Siberian haplotypes (fig. 6). This phylogeny shows that all Asian mtDNAs can be subdivided into two macroclusters defined by the presence or absence of the polymorphic site at *Dde*I at np 10394, which also bifurcates the European mtDNA lineages. Moreover, every Asian mtDNA that harbors the *Dde*I site at np 10394 also has an adjacent *Alu*I site at np 10397 (C to T at np 10400). The constant association of the *Dde*I np 10394 and *Alu*I np 10394 in Asians, but not in Africans or Europeans, implies that the *Alu*I np 10397 mutation must have arisen on an mtDNA carrying the *Dde*I np 10394 mutation shortly after women arrived in Asia.

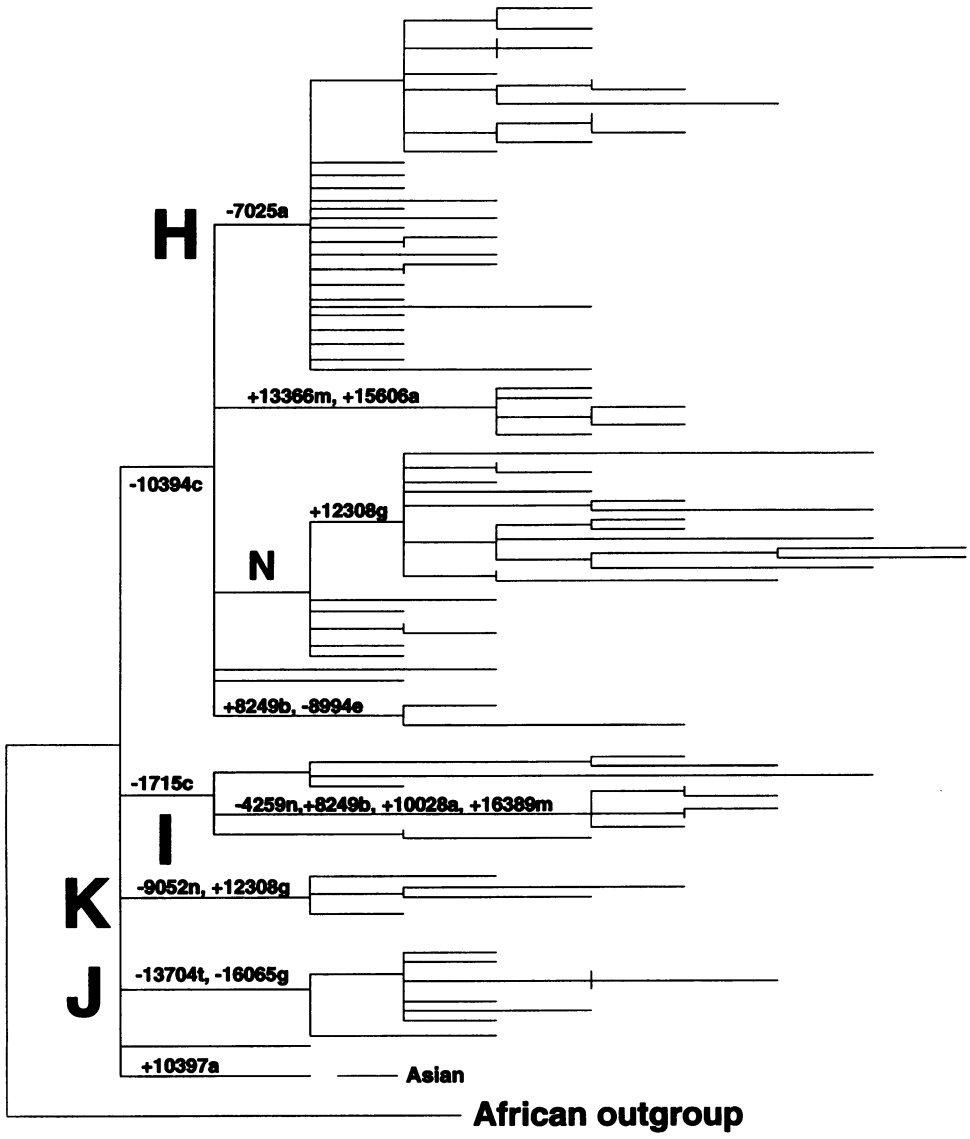

Figure 5 Representative maximum parsimony tree of European-derived mtDNAs. The more informative restriction-site polymorphisms that define the major branches (capital letters) are designated according to the nomenclature of fig. 3. The data and phylogenic analysis from which this tree was derived are presented in the work of Torroni et al. (1994*b*).

In addition to this major subdivision of Asian mtDNAs, there are a number of distinctive sublineages of relevance to Asian and Native American prehistory. Haplogroups A, B, C, and D have proved to be the progenitors of virtually all Native American mtDNAs. Haplogroups A and B lack the *Dde*I site at np 10394 and the *Alu*I site at np 10397, while haplogroups C and D have these sites. In addition, haplogroup A is defined by a *Hae*III site at np 663 (A to G at np 663), haplogroup B by an independent occurrence of the 9-np deletion between the COII and $tRNA^{Lys}$ genes, haplogroup C by the simultaneous *Hinc*II site loss at np 13259 and an *Alu*I site gain at np 13262 (A to G at np 13262), and haplogroup D by the loss of an *Alu*I site at np 5176 (C to A at np 5178). These haplogroups are further delineated in most Asians and Native Americans by specific control-region variants. For haplogroup A, these include variants at nps 16362 (T → C), 16319 (G → A), 16290 (C → T), and 16223 (C → T); for haplogroup B, variants at nps 16217 (T → C) and 16189 (T → C); for haplogroup C, variants at nps 16327 (C → T), 16298 (T → C), and 16223 (C → T); and, for haplogroup D, variants at nps 16362 (T → C) and 16223 (C → T) (Torroni et al. 1993*a*).

Three other prominent Asian haplogroups are E, F, and G. Haplogroups E and G have the combined *Dde*I and *Alu*I sites at nps 10394 and 10397, while haplogroup F lacks these sites. Haplogroup E is further de-

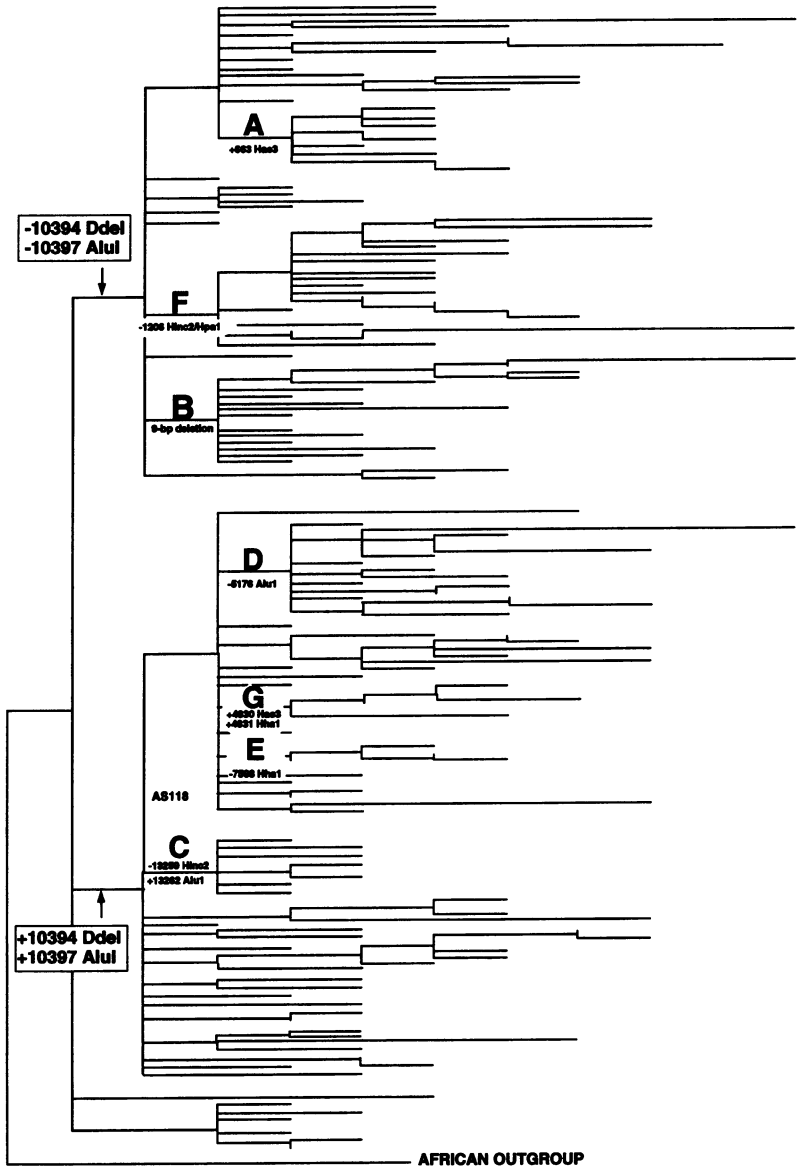

Figure 6 Representative maximum parsimony tree of Asian mtDNAs. The more informative restriction sites that define the major branches (capital letters) are designated according to the nomenclature of fig. 3. The data and phylogenetic analysis from which this tree were derived are presented in the work of Torroni et al. (1994*c*).

fined by a *Hha*I site loss at np 7598, Haplogroup G by the presence of a *Hae*III site at np 4830 and a *Hha*I site at np 4831, while haplogroup F is delineated by the combined *Hpa*I/*Hinc*II site loss at np 12406, the first Asian-specific polymorphism observed (Denaro et al. 1981; Blanc et al. 1983).

All of these haplogroups show marked frequency variation throughout Asia. Haplogroup F is prominent in southern Asian populations, being found in 32% of Vietnamese mtDNAs and 21% of Malay mtDNAs. It is present in ~15% of Koreans and Tibetans but is virtually absent in Siberia. By contrast, haplogroups A, C, D, E, and G are absent in southern Asian populations, including Vietnamese, Malays, Sabah, Malay aboriginals, and New Guineans but are found at significant frequencies in Tibetans and Koreans or Han Chinese. This north-south distinction supports the dichotomization of Asians into the Sinodont (northern) and Sunodont (southern) Asian populations (Turner 1983, 1987). Furthermore, haplogroups A, C, and D extend into Siberia. Haplogroups C and D are prevalent in most of the nine Siberian populations analyzed, reaching maximum frequencies of 84% and 34%, respectively. Haplogroup A reaches its highest frequencies in the Chukchi and

Koryaks, the northeasternmost populations of Siberia and likely progenitors of Native Americans. The haplogroup frequencies of the Koryaks are 24% A, 22% C, 9% D, and 46% other, while those of the Chukchi are 38% A, 17% C, 17% D, and 29% other (Torroni et al. 1993*b*).

Haplogroup B, defined by the 9-np COII-tRNA^{Lys} deletion, displays a markedly different distribution. It is common throughout central and southern Asia and is prominent in coastal Asian populations, approaching fixation (100%) in certain Pacific island populations (Hertzberg et al. 1989; Stoneking et al. 1990; Ballinger et al. 1992*a*). It is completely absent from all nine Siberian populations analyzed yet reappears throughout North, Central, and South American Native American populations (Schurr et al. 1990; Torroni et al. 1993*a*, 1993*b*). The high frequency of this haplogroup among coastal Asian and Pacific island populations and its striking absence in Siberians relative to Central Asians and Native Americans suggest that haplogroup B mtDNAs did not come to the Americas via a trans-Siberian migration, but rather may have crossed from Asia to the Americas by migration along the Siberian coast.

The co-occurrence of the *Alu*I site at np 10397 and the *Dde*I site at np 10394 in all Asian mtDNAs indicates that the *Alu*I site gain occurred at the beginning of Asian habitation. Consequently, the sequence diversity that has accumulated in the *Dde*I np 10394–*Alu*I np 10397 lineage should be indicative of the age of the Asian population. The overall sequence diversity in this lineage is 0.161%. This gives an age for the Asian population of 56,000–73,000 YBP (fig. 4).

Native American mtDNA variation.—To learn more about the origin of Native Americans, we analyzed 743 Native American mtDNAs. Multiple hypotheses have been put forward to explain the origin and radiation of Native Americans. One hypotheses is based on the classification of Native American languages by Greenberg et al. (1986). Greenberg et al. divided all Native American languages into three major groups: Amerind, which encompasses the great diversity of languages spoken by the Paleo-Indian peoples occupying most of North America and all of Central and South America; Na-Dené, which is spoken by the Athapaskans of the northeastern United States, Canada, and Alaska as well as the Navajo and Apache who migrated south through the great plains around 1,000 B.C.; and the Eskaleut languages, which are spoken by the Eskimos and Aleuts of the Arctic region. Greenberg et al. hypothesized that each of these language groups corresponded to a different migration, arising in a different geographic homeland. Using glottochronology—dating based on the divergence rate of languages—they estimated that these migrations occurred at ≥11,000 YBP, 9,000 YBP, and 5,000 YBP, respectively.

In our first studies on Native American mtDNA variation, we focused on the Pima and Papago, Paleo-Indians of the southwestern United States. Using Southern blot analysis and our initial six informative restriction endonucleases, we discovered that ~40% of these Native American mtDNAs lacked the *Hinc*II site at np 13259 (Wallace et al. 1985), while only 1.8% of central Asian mtDNAs lacked this site (Blanc et al. 1983). This led to the hypothesis that Native American mtDNAs were derived from a limited number of founding mtDNA haplotypes that crossed the Bering land bridge in distinct migrations (Wallace et al. 1985; Schurr et al. 1990; Wallace and Torroni 1992). This hypotheses has subsequently been confirmed by our more extensive analysis encompassing 563 Paleo-Indians from 24 tribes, 130 Na-Dené representing five tribes, and 50 Eskimos (Torroni et al. 1992, 1993*a*, 1994*a*, 1994*d*).

Our analysis of *mtDNA variation in Paleo-Indians* revealed a dramatic result. Virtually all of the mtDNAs fell into one of the four Asian haplogroups: A (*Hae*III site at np 663), B (9-np deletion between COII/tRNA^{Lys}), C (*Hinc*II site loss at np 13259 and *Alu*I site gain at np 13262), and D (*Alu*I site loss at np 5176), with haplogroups C and D also harboring the *Dde*I np 10394 and *Alu*I np 10397 site gains found in their Asian progenitors (Schurr et al. 1990; Torroni et al. 1992, 1993*a*) (fig. 7). Each of the four Native American haplogroups traces back to a single model mtDNA haplotype that is shared by Asia and the Americas and that initiated the mtDNA radiation in the Americas. However, none of the derived haplotypes are shared by Asians and Americans, as demonstrated by analyzing haplogroup C and D mtDNAs from Siberians and Native Americans (Torroni et al. 1993*b*). Hence, it appears that only four mtDNA haplotypes crossed from Siberia into the Americas. All four of these haplogroups are distributed throughout the Paleo-Indians of North, Central, and South America, though individual tribes may have lost one or more of the haplogroups through genetic drift. The broad distribution of all four mtDNA haplogroups suggests that they either came together or were subsequently thoroughly mixed.

The calculation of the mtDNA sequence diversity that has accumulated within each haplogroup revealed that haplogroups A, C, and D had relatively similar values: A = 0.075%, C = 0.096%, and D = 0.053%, with a mean value of 0.075%. By contrast, haplogroup B had a much lower value, 0.034%, suggesting that haplogroup B arrived in the Americans much more recently than A, C, and D (Chen et al. 1995). This difference is consistent with the absence of haplogroup B in Siberia, even though haplogroups A, C, and D are prevalent. These two results imply that the Paleo-Indians of the Amerind linguistic group may have been derived from two migrations. The first migration moved up from central Asia through Siberia, during which it became pro-

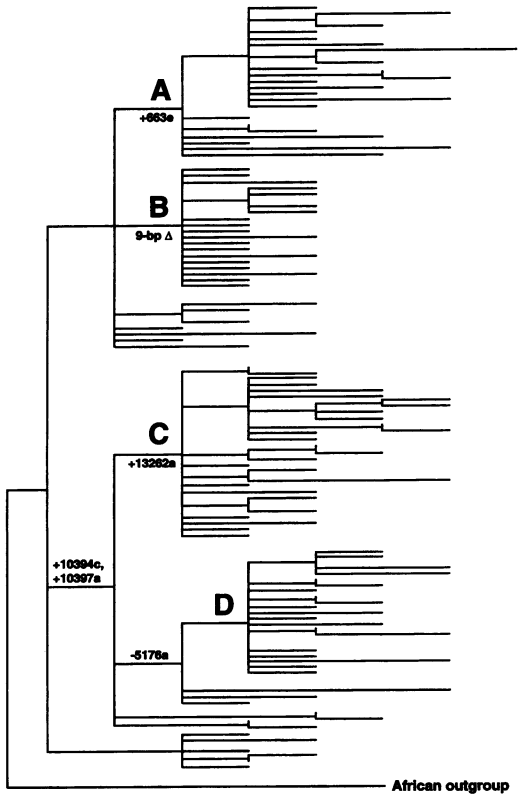

Figure 7 Representative maximum parsimony tree of Native American mtDNAs. The more informative restriction sites that define the major branches (capital letters) are designated according to the nomenclature of fig. 3. The data and phylogenetic analysis from which this tree were derived are provided in the work of Torroni et al. (1993*a* and 1994*d*) and of Torroni and Wallace (1995).

gressively enriched for the founder haplotypes of haplogroups A, C, and D. Ultimately, only these haplotypes crossed the Bering land bridge to found the Paleo-Indians. The second migration came much later, bearing the founder haplotype of haplogroup B. This migration bypassed Siberia, possibly moving along the Siberian and Alaskan coasts, and entered the Americas. The B migration must have encountered relatively low-density populations harboring haplogroups A, C, and D and dispersed throughout their range to give the modern Paleo-Indians (Schurr et al. 1990; Torroni et al. 1992, 1993*a*).

To further investigate the Paleo-Indian tribalization process, we examined the mtDNAs from an isolated group of Aztec descendants in central Mexico: the Mixtec of Alta and Baja, the Zapotecs, and the adjacent Mixe. These tribes were compared with the Pima of Arizona, the Maya of Yucatan, the Chibchan-speakers of Panama, the Bella Coola and the Nuu-Chah-Nulth of North America, and the Yanomama and Wapishana of South America. In aggregate, the Aztec-derived Mixtec and Zapotecs harbored three haplogroups: 66% A, 18% B, and 16% C. The linguistically related Pima of Arizona harbored the same haplogroups, as did the Mixe, suggesting a common ancestry. The Maya were more similar to the North American Paleo-Indians, while the Chibchans and South American tribes differed from the more northern tribes and from each other. These results suggest that the Maya and Aztecs may have been derived from different populations (Torroni et al. 1994*a*).

The prevalence of haplogroup A, B, C, and D mtDNAs in Native American populations has now been confirmed by multiple investigators (Ward et al. 1991, 1993; Horai et al. 1993; Santos and Barrantes 1994). These same four haplogroups have also been found in the Native American skeletons excavated from a pre-Columbian burial site (Stone and Stoneking 1993). While some Native American mtDNAs have been found to not exhibit one of the four primary mtDNA restriction site markers (Torroni et al. 1992, 1993*a*; Bailliet et al. 1994), most of these mtDNAs can be shown to result from either recent genetic admixture with European or African immigrants or the secondary gain or loss of informative restriction sites (Torroni and Wallace 1995).

This is most evident for the Ojibwa of the Great Lakes region. Of 42 Ojibwa samples analyzed 11 (26%) were not from haplogroups A, B, C, or D (non-A–D). Similarly, of the Nuu-Chah-Nulth and the Bella Coola of the Pacific northwest, 13.3% and 4%, respectively, were non-A–D. The most likely source of the non-A–D haplotypes is admixture with European populations. Studies of blood group markers have suggested that the Ojibwa have experience 30% European admixture (Szathmary 1984). Complete haplotype analyses of seven Ojibwa, as well as several other "anomalous" cases, revealed that they lacked a *Dde*I site at np 1715 and a *Dde*I site at np 10394 (Torroni et al. 1993*a*). This unique haplotype was not observed in all 411 Siberians (Torroni et al. 1993*b*) or 207 Asians (Ballinger et al. 1992*a*; Torroni et al. 1994*c*), indicating that it was probably not of Asian ancestry. On the other hand, in our survey of 175 European mtDNAs, 2 were found to have similar haplotypes to the "anomalous" Ojibwa mtDNAs (Torroni et al. 1994*b*); and in a recent survey of Finnish mtDNAs two additional mtDNAs were found with the same haplotype as seen in the Ojibwa. Hence, it now appears that most, if not all, Paleo-Indians mtDNAs were derived from only four founding mtDNA haplotypes, with additional mtDNAs being added through recent admixture.

Because the early immigrants from Europe to the Americas were predominantly male, isolated Native American populations might be expected to contain exclusively Native American mtDNAs, while still having acquired significant European nuclear genes via males.

This possibility is being tested using continent-specific Y chromosome variation. We have analyzed the Y chromosomes of the Mixtec, Zapotec, and Mixe tribes of central Mexico, populations which contain only mtDNAs from haplogroups A, B, and C, indicating exclusively Native American female ancestry. By contrast, the Y chromosome polymorphisms suggests that these tribes have experienced non–Native American male admixture.

Like the mtDNA, the Y chromosome has been found to harbor continent-specific polymorphisms. The most extensively studied polymorphic system is that of the DYS1 locus located on the long arm at Yq11.2 (Ngo et al. 1986). Polymorphisms at this locus are detected by Southern blot of *Taq*I-digested DNA, hybridized to cosmid 49 fragments "a" and "f." This gives an array of ~18 fragments, several of which are variable in their presence and/or length. The array of fragments observed for a Y chromosome is defined as a haplotype (Ht) with >70 haplotypes having been defined for Europeans and Africans. In Europeans, the most common haplotypes are Ht15, Ht12, Ht24, and Ht28, which encompass 40%–60% European males; while in Africans haplotypes Ht4, Ht5, Ht10, Ht23, Ht30, and Ht33 account for 75%–100% of Y chromosomes (Lucotte et al. 1989; Torroni et al. 1990; Spurdle and Jenkins 1991; Persichetti et al. 1992; Santachiara-Benerecetti et al. 1993). While Asian and Native American Y chromosomes have not been explored, analysis of 31 Y chromosomes from the Mixtec, Zapotec, and Mixe revealed that 45% were Ht18, 13% were Ht13, and 13% were Ht63. These haplotypes are relatively rare in Europe and Africa and probably represent the common founding Native American Y chromosome haplotypes. In addition, two of the Y chromosomes harbored the European variant Ht15 and one Ht1. Moreover, one Y chromosome had the African haplotype Ht5 (Torroni et al. 1994*a*). Thus, between 2 and 4 of the 31 Y chromosomes (6%–13%) may have been the result of European and African admixture. While more extensive analysis of Asian and Native American Y chromosomes will be necessary to confirm that Ht15, Ht1, and Ht5 were not present in the Native American progenitors, the current data suggest that there has been some European and African male admixture in these isolated populations.

Analysis of *mtDNA variation of the Na-Dené* has confirmed that they are distinct from the Paleo-Indians. Analysis of northern Na-Dené, including Dogrib, Tlingit, and Haida, indicates that these people harbor mtDNAs from only one of the haplogroups, haplogroup A (Torroni et al. 1992, 1993*a*). This is substantiated by the southern Na-Dené, including the Apache and Navajo, who are >60% haplogroup A. The remaining Apache and Navajo mtDNAs are from haplogroups B, C, and D and probably represent recent admixture with adjacent Paleo-Indian tribes. The distinctive nature of the Na-Dené mtDNAs is confirmed by the fact that about a third of all Na-Dené haplogroup A mtDNAs carry a novel variant, an *Rsa*I site loss at np 16329 (A→G at 16331). Interestingly, this variant is found in all Na-Dené but the Haida. This may suggest that the *Rsa*I site loss occurred during the migration of the Na-Dené from Siberia to the Americas, such that certain populations are derived only from the founder haplotype.

Analyses of the sequence diversity of the Na-Dené haplogroup A mtDNAs gave a value of 0.021%. This is substantially lower than the diversity of the Paleo-Indian haplogroups A + C + D and also different from the Paleo-Indian haplogroup B. Hence, the Na-Dené do appear to have arrived as a single migration, which occurred more recently than those of the Paleo-Indians, consistent with conjectures from linguistic data (Torroni et al. 1992, 1993*a*).

Analysis of the mtDNA variation in Eskimos has been more difficult to study, because of the limited availability of samples and their recent divergence from ancestral populations. However, a survey of 50 Siberian Yupik Eskimos, who represent Eskimo peoples inhabiting both sides of the Bering Strait, revealed only haplogroups A and D. Hence, the Eskimos may also be distinct from the Na-Dené and Paleo-Indians, as predicted from linguistic associations (Torroni et al. 1993*b*).

Times of the Native American migrations, as in the cases of the African, European, and Asian populations, were calculated using the haplogroup-specific mtDNA sequence diversity and the mtDNA sequence evolution rate of 2.2%–2.9%/MYR. This rate was calculated using the Chibcha-speaking peoples of Central America, who are estimated from anthropological and nuclear genetic data to have originated ~8,000–10,000 YBP. Analyzing the mtDNA variation of 110 Chibchans representing five tribes revealed that all but one of the Chibchan mtDNAs were from haplogroups A or B, the exception being a single haplogroup D. Moreover, 62% of the Chibchan haplogroup A mtDNAs showed a distinctive private polymorphism, the loss of an *Msp*I site at np 104, the result of a small deletion in the 3′ end of the D-loop hypervariable region II (Santos and Barrantes 1994; Torroni et al. 1994*d*). Averaging the sequence diversity of the haplogroup A and B mtDNAs and dividing by the putative age of the population gave the sequence evolution rate of 2.2%–2.9%/MYR (Torroni et al. 1994*d*).

Using this Native American sequence evolution rate, we calculated that the first Paleo-Indian migration carrying haplogroups A, C, and D (sequence diversity of 0.075%) arrived 26,000–34,000 YBP and that the second Paleo-Indian migration bringing haplogroup B (0.034%) arrived 12,000–15,000 YBP. The Na-Dené

migration with a sequence diversity of 0.021% was estimated to have arrived 7,200–9,000 YBP, a value strikingly similar to the 9,500 YBP value estimated by glottochronology (Torroni et al. 1994*d*) (fig. 4).

mtDNA Mutations and Disease

The assessment of global mtDNA variation has not only been valuable for reconstructing human origins, it has also been essential for elucidating the role of mtDNA mutations in disease. This is true for two reasons. First, it is necessary to know what the nonpathogenic mtDNA variation is so that it can be distinguished from pathogenic mutations. Second, mildly deleterious mutations that cause symptoms in the postreproductive period may only minimally reduce reproductive fitness and hence may have characteristics in common with neutral polymorphisms.

Since mutations are essentially random, all possible mtDNA base substitutions must arise in the human female germ line. The neutral mutations segregate to homoplasmy and ultimately become established in the population by genetic drift, a process taking thousands of years. Consequently, neutral polymorphisms are almost always homoplasmic and are consistently associated with the background haplotype on which they arose. Similarly, mildly deleterious mutations that cause disease late in the life may also become established in the population as low-frequency polymorphisms, be homoplasmic, and consistently be found on the same background haplotype. By contrast, deleterious mutations cause earlier-onset disease as they segregate toward homoplasmy by replicative segregation. Hence, these are rapidly eliminated by natural selection, so that currently existing mutations must represent new independent mutations, each associated with a different background mtDNA haplotype (Wallace 1994).

Pathological mtDNA mutations can be divided into two classes: missense mutations that alter polypeptide genes and protein synthesis mutations that alter rRNAs and tRNAs genes (Wallace 1992*a*, 1994; Wallace et al. 1995). The best-studied clinical phenotypes associated with missense mutations are Leber hereditary optic neuropathy (LHON) and neurogenic muscle weakness ataxia and retinitis pigmentosa (NARP) in association with Leigh syndrome.

LHON and dystonia.—LHON is a form of midlife onset, acute or subacute blindness, leading to central scotoma. In familial cases, all affected individuals are related through the female lineage. However, for reasons that are unclear, only a portion of the maternal relatives loose their vision (Wallace and Lott 1992; Shoffner and Wallace 1994). While 19 mutations have been reported to be associated with LHON, 5 appear to play a primary role in causing the disease (Brown and Wallace 1994) (fig. 1). These mutations are a G-to-A transition at np 14459 in the ND6 gene (MTND6*LDYT14459A) (Jun et al. 1994), a G-to-A transition at np 11778 in the ND4 gene (MTND4*LHON11778A) (Wallace et al. 1988*a*), a G-to-A transition at np 3460 in the ND1 gene (MTND1*LHON3460A) (Howell et al. 1991; Huoponen et al. 1991), a T-to-C transition at np 14484 in the ND6 gene (MTND6*LHON14484C) (Johns et al. 1992), and a G-to-A transition at np 15257 in the cytb gene (MTCYB*LHON15257A) (Johns and Neufeld 1991; Brown et al. 1992*a*). While all five of these mutations can result in the same phenotype, they differ in their severity, as revealed by their ability to cause additional neurological symptoms, propensity for causing the disease in the absence of other contributing genetic factors, the frequency with which they are heteroplasmic, and their tendency for visual recovery.

The MTND6*LDYT14459A mutation is by far the most severe of the LHON mutations and provides a good example of the utility of knowing the continent-specific variation of the mtDNA when attempting to identify pathogenic mtDNA mutations. This mutation was first discovered in a large Hispanic pedigree in which two very different clinical phenotypes were observed in maternal relatives. One phenotype was LHON, but the other involved more severe neurological symptoms including early-onset generalized dystonia (a movement disorder involving progressive rigidity associated with basal ganglia degeneration or bilateral striatal necrosis) together with pseudobulbar syndrome, short stature, and reduced intelligence. Among the 42 maternal relatives in this family, 19% had LHON, 31% dystonia, and 2% both (Novotny et al. 1986). Sequence analysis of this family's mtDNA revealed 40 nucleotide differences relative to the reference European sequence. Since several of these were potentially pathogenic, it was initially impossible to identify the causal mutation. This dilemma was resolved when we began to characterize Native American mtDNA variation and discovered that the LHON-dystonia family harbored an mtDNA from Native American haplogroup D. By testing each of the base substitutions from the patient's mtDNA against a collection of Native American haplogroup D mtDNAs, all but one of the base substitutions potentially affecting gene function could be shown to be haplogroup-specific polymorphisms. The exception was a mutation in ND6 at np 14459. This mutation converted a relatively highly conserved alanine to a valine in the ND6 gene and proved to be heteroplasmic (Jun et al. 1994). A subsequent survey of dystonia and LHON patients for this mutation revealed two additional positive families: an isolated Caucasian child with dystonia and an European mtDNA haplotype and an African-American mother and daughter with LHON and an African mtDNA haplotype, with both families being heteroplasmic (Shoffner et al., in press). Thus, the np 14459 mutation has arisen

in each of these families independently and while still heteroplasmic can result in either LHON or dystonia.

While the MTND6*LDYT14459A mutation is the most severe LHON mutation, the MTND4*LHON-11778A mutation is the most common cause of LHON, accounting for 50% of European cases and >90% of Asian cases. MTND4*LHON11778A is also the next most severe LHON mutation, changing a highly conserved arginine to a histidine. This mutation is sufficient in itself to cause LHON, is heteroplasmic in ~14% of families, and shows spontaneous visual recovery in only ~4% of cases. MTND1*LHON3460A is the next-most-severe mutation. It accounts for ~15% of European cases, changes a moderately conserved alanine to a threonine, is occasionally heteroplasmic, and shows spontaneous visual recovery in ~22% of cases. MTND6*LHON14484C is a significantly milder mutation. It accounts for ~15% of cases, changes a weakly conserved methionine to a valine, is very rarely heteroplasmic, is associated with 37% spontaneous visual recovery, and is frequently found in association with other LHON mutations, particularly MTCYB*LHON-15257A. Finally, MTCYB*LHON15257A is the mildest of the LHON mutations. It is found in ~9% of European cases, changes a highly conserved aspartate to an asparagine, is always homoplasmic, is associated with ~28% visual recovery, and is frequently found together with other LHON mutants, particularly MTND6*LHON14484C (Wallace 1994).

The relative pathogenicity of these LHON mutations can be further assessed by phylogenetic analysis (Brown et al., in press). The more severe mutations, which significantly reduce reproductive fitness, will be represented primarily by new mutations dispersed among different mtDNA haplotypes. By contrast, milder mutations, which do not significantly reduce reproductive fitness, will have a common origin and a shared haplogroup. Haplotype analysis of patients with the MTND4*-LHON 11778A and MTND1*LHON3460A mutations have revealed that each separate family has a different background haplotype. Hence, virtually every case must be a new mutation, and these mutations must be relatively highly pathogenic. By contrast, the haplotypes of patients harboring the MTND6*LHON14484C mutation tend to cluster in the J haplogroup, with only two independent cases occurring in a second European lineage. Moreover, both of these lineages harbor secondary LHON mutations that could augment the pathogenicity of the MTND6*LHON14484C mutation. The J haplogroup has the MTND5*LHON13708A, MTND1*-LHON3394C, and MTCYB*LHON 15257A variants (Brown et al. 1992*a*, 1992*b*, and in press), while the second lineage harbors the MTCOI*LHON7444A mutation (Brown et al. 1992*c*). Similarly, the MTCYB*-LHON15257A mutation is exclusively confined to the J haplogroup (Brown et al. 1992*b*, and in press). The frequent association of the MTCYB*LHON15257A and MTND6*LHON14484C mutations with specific background haplotypes indicates either that they require additional mtDNA variants to augment their pathogenicity and expression or that they have arisen a limited number of times and persisted in the population as low-frequency polymorphisms. In either case, the MTND6*-LHON14484C and MTCYB*LHON15257A mutations must be significantly less pathogenic than the MTND4*LHON11778A and MTND1*LHON3460A mutations (Brown et al., in press).

Knowledge of continent-specific mtDNA variation has also been important for demonstrating that certain LHON-like clinical presentations are not the product of mtDNA mutations. In Cuba, there has been a recent epidemic of optic neuropathy that has blinded 26,000 individuals. The Cuban government requested that we determine whether this outbreak was due to an mtDNA mutation. A clinical survey was conducted of patients and normals in the rural Pinar del Rio province, and 57 patients and 69 controls were tested for the eight most common LHON-associated mutations. None of the mutations was either common in the patients or at a higher frequency in patients than in the controls, thus ruling out the known mtDNA LHON mutations (Newman et al. 1994).

However, it was still possible that the affected Cubans harbored a new mtDNA mutation or a continent-specific polymorphism that predisposed them to blindness when exposed to an environmental stress. If either of these possibilities were the cause of the epidemic blindness, then the affected Cubans should share a common or similar mtDNA haplotype, the haplotype on which the new mutation arose or which carried the responsible polymorphism. Hence, the patients would have much more uniform mtDNA genotypes than the surrounding population.

The current Cuban population represents an amalgamation of Native American, European, and African populations. Hence, Cubans harbor a wide variety of mtDNA haplotypes. A survey was conducted of the Pinar del Rio samples for the presence or absence of 21 polymorphic mtDNA markers, which in aggregate could identify nearly 100% of the Native American mtDNAs, 68% of European mtDNAs, and 70% of African mtDNAs. This revealed that 3.7% of the Cuban mtDNAs were of Native American origin and included haplogroups A, B, and D; that 26% were European including haplogroups H, I, J, M, and O; and that 31% were African from haplogroup L. Knowing the continental frequencies of these European and African haplogroups, we could calculate the matritineal input from these continents into Cuba. This revealed that Africans contributed 46% of the mtDNAs, Europeans contrib-

uted 50%, and Native Americans contributed 4% (Torroni et al. 1995).

Comparing the haplotype distributions of the patients and the controls revealed that both groups had similar haplogroup distributions. Since the patients showed no bias toward a particular haplogroup, the Cuban epidemic optic neuropathy is unlikely to have been the result of a continent-specific polymorphism or a new Cuban LHON mutation.

NARP and Leigh syndrome.—Two mutations have been identified in the mtDNA ATP6 gene that cause neurodegenerative disease. Both occur in codon 156 and are highly pathogenic, suggesting that most if not all independent families are independent mutations. The MTATP6*NARP8993G mutation converts the highly conserved leucine to an arginine (Holt et al. 1990), while the MTATP6*NARP8993C converts the leucine to a proline (De Vries et al. 1993). The two mutations are invariably heteroplasmic, and as they segregate they generate a wide range of neurological symptoms from mild retinitus pigmentosa through macular degeneration, mental retardation, and olivopontocerebellar atrophy to Leigh syndrome (Shoffner et al. 1992; Tatuch et al. 1992; Ortiz et al. 1993). Leigh syndrome is a frequently lethal childhood disease associated with the progressive degeneration of the basal ganglia. Presumably, the variable clinical presentation of np 8993 mutations reflects the replicative segregation of the heteroplasmic mutations.

The biochemical abnormality of the MTATP6*-NARP8993G mutation was shown to be a defect in the proton channel of the ATP synthase. Lymphoblastoid cell lines harboring the mutation were prepared from patients, and the mitochondria were analyzed for OXPHOS defects by respiration studies. Respiratory complexes I, III, and IV of the electron transport chain were shown to be normal, since the patient's mitochondria had the same maximum respiration rate as normal mitochondria, when respiring in the presence of an uncoupler nitrophenol. However, respiration coupled to ATP synthesis and the ADP/0 ratio were both reduced 30%–40% in the patient mitochondria, suggesting a defect in the proton channel of the ATP synthase. These defects were then linked to the MTATP6*8993G mutation using cybrids transfer. Heteroplasmic patient lymphoblasts were enucleated, and the cytoplasmic fragments were fused to recipient cells that had previously been cured of their mtDNAs (ρ° cells) (King and Attardi 1988). The resulting cybrids proved to be either homoplasmic mutant, homoplasmic normal, or heteroplasmic. Respiration studies on mitochondria isolated from homoplasmic normal cybrids were normal, while mitochondria from homoplasmic mutant cybrids showed the same 30%–40% reduction in ADP-stimulated respiration and ADP/0 ratio as the parental patient cell lines. Hence, the ATP synthase defect and the variable clinical phenotypes are both the result of the heteroplasmic T-to-G transversion at np 8993 in the ATP6 gene (Trounce et al. 1994).

Pathological mtDNA Protein Synthesis Mutations

Close to 30 mtDNA tRNA and rRNA gene mutations have been associated with various degenerative diseases (Wallace et al. 1995). These range in severity from lethal to mild. Since moderate to severe protein synthesis mutations cause severe clinical symptoms in children and young adults, they substantially reduce reproductive fitness. Hence, these mutations are routinely heteroplasmic, with each family harboring a different mutation. By contrast, very mild protein synthesis mutations may not cause symptoms until well after reproductive age. These mutations are frequently homoplasmic and can be maintained as low-frequently polymorphisms in the general population for many generations (Wallace 1994).

Eleven of the tRNA mutations are located in only one tRNA, $tRNA^{Leu(UUR)}$ (Wallace et al. 1995). The clinical phenotype's associated with these mutations range from the lethal progressive encephalomyopathy (PEM) associated with a single base deletion at np 3271; through pediatric and adult hypertrophic cardiomyopathy and myopathy; to the mitochondrial myopathy, encephalomyopathy, lactic acidosis, and stroke-like episodes syndrome (MELAS) and type II diabetes and deafness, both associated with an A-to-G transition at np 3243.

The MTTL1*PEM3271 mutation was discovered in a single individual who died of tonic colonic seizures, renal failure, and sepsis at the age of 28 years (Shoffner et al. 1995) (fig. 1). Clinical symptoms first appeared at age 5 years, with progressive hearing loss, resulting in deafness by age 18 years. Seizures also occurred during childhood, with progressive neuromuscular degeneration occurring as a young adult and culminating in mitochondrial myopathy associated with ragged-red muscle fibers and abnormal mitochondria, retinitus pigmentosa, glaucoma, dementia, and severe cerebral calcifications (Fahr disease). The mutation resulted in the deletion of one of three Ts, which make up the stem of the anticodon loop, arbitrarily designated np 3271. The patient was heteroplasmic, though predominantly mutant, while the mother was homoplasmic normal. Hence, this mutation segregated from homoplasmic normal to virtually homoplasmic mutant in one generation (Shoffner et al. 1995). In fact, preliminary observations suggest that there may be a trend in which the more severe mutations segregate more rapidly to homoplasmy and thus are associated with smaller families.

The milder MTTL1*MELAS3243G mutation is frequently associated with larger heteroplasmic pedigrees (fig. 1). Individuals harboring high percentages of mu-

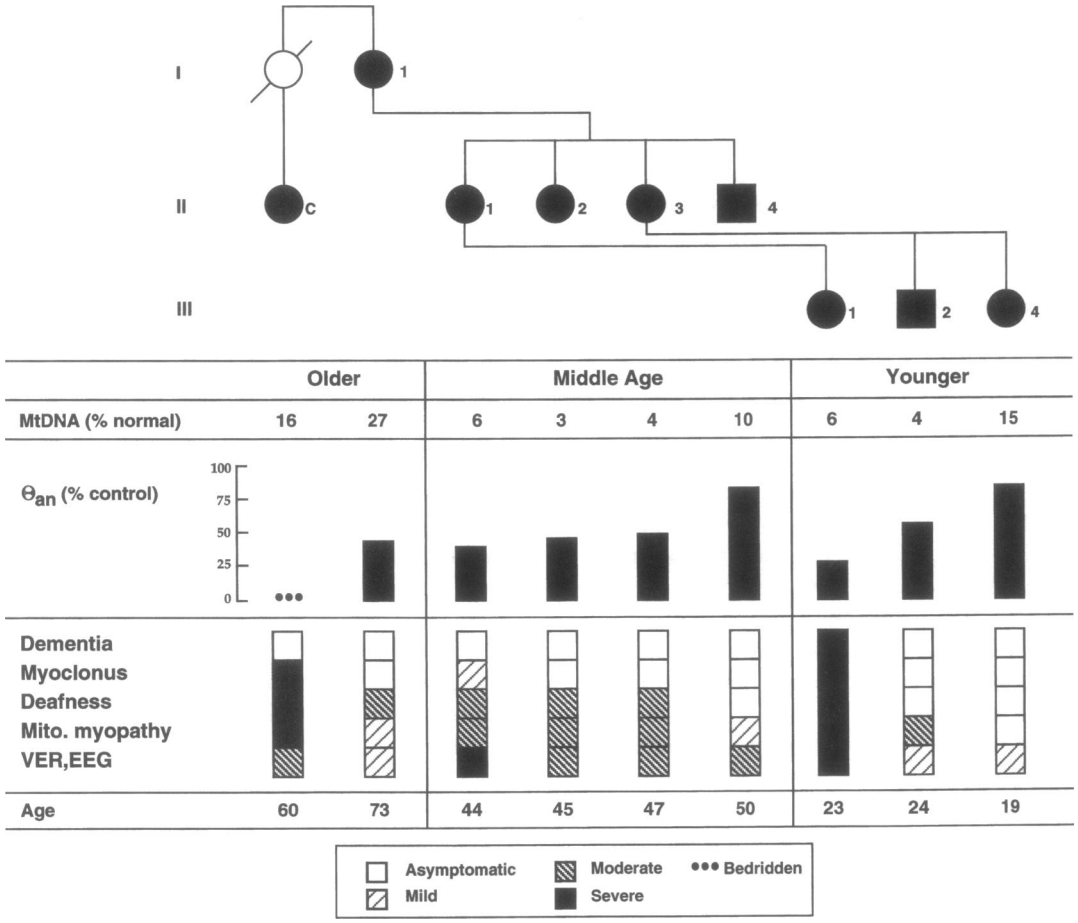

Figure 8 Three-generation MTTK*MERRF8344G pedigree, showing the interrelationship between phenotype and the combination of mtDNA genotype and age. θ_{an} = anaerobic threshold, an indicator of the skeletal muscle OXPHOS capacity; Mito. = mitochondrial; VER = visual evoked response abnormalities; and EEG = electroencephalograph abnormalities. Reprinted, with permission, from the article by Wallace et al. (1994).

tant mtDNAs can experience severe, even lethal, symptoms, including mitochondrial myopathy, stroke-like episodes, hypertrophic cardiomyopathy, and renal failure (Goto et al. 1990; Kobayashi et al. 1990). However, families and individuals with lower percentages of mutant mtDNA can present with only adult-onset diabetes mellitus and neurosensory hearing loss (van den Ouweland et al. 1992, 1994). In fact, between 1%–3% of all type II diabetes are the result of the MTTL1*-MELAS3243G mutation (Kadowaki et al. 1994).

A common clinical feature of protein synthesis mutations is the delayed onset of symptoms, followed by clinical progression. This feature has been best documented for the myoclonic epilepsy and ragged-red fiber disease syndrome (MERRF) (fig. 1) (MTTK*MERRF-8344G) (fig. 1) (Wallace et al. 1988*b*, 1994; Shoffner et al. 1990). Pedigrees harboring this mutation are frequently large, invariably heteroplasmic, and clinically variable. In one well-characterized family (fig. 8), the clinical manifestations of maternal relatives ranged from essentially normal through mitochondrial myopathy and hearing loss to severe multisystem degenerative disease including incapacitating myoclonic jerking, mitochondrial myopathy, hearing loss, dementia, renal dysfunction, and cardiomyopathy. All maternal relatives were heteroplasmic for the mutation, with the clinical symptoms being associated with two variables: the percentage of mutant mtDNAs the individual inherited at birth and the individual's age. For example, a 20-year-old with 85% mutant mtDNAs was normal, while a comparably aged cousin with 96% mutant mtDNAs was severely affected. Moreover, a 60-year-old women with 84% mutant mtDNAs was severely affected with myoclonus, mitochondrial myopathy, deafness, and migraines, while a 73-year-old family member with 73% mutant mtDNAs had mild mitochondrial myopathy and hearing loss. These results imply that the inherited mtDNA defect defines the range of clinical symptoms that an individual will experience and the approximate age at onset but that a secondary aging effect is responsible for the subsequent

progression and final outcome of the disease (Wallace et al. 1988*b*, 1994; Shoffner et al. 1990) (fig. 8).

The mildest tRNA mutation described to date is the MTTQ*ADPD4336C mutation associated with late-onset Alzheimer disease (fig. 1). This mutation changes a mildly conserved A to a G in the $tRNA^{Gln}$ gene between the amino acid acceptor stem and the TΨC stem. The mutation is found in 5% of late-onset Alzheimer patients but in only 0.4%–0.7% of the general population (Shoffner et al. 1993; Cortopassi and Hutchin 1994). Moreover, all individuals with the mutation have the same background haplogroup H, indicating that this mutation arose as a single event 15,000–20,000 YBP and created a maternal lineage predisposed to Alzheimer disease. In addition to the MTTQ*ADPD4336C mutation, Alzheimer disease patients of this lineage can also harbor additional potentially pathogenic mutations. These include an MTND1*ADPD3397G mutation that changes a highly conserved methionine to a valine and an insertion of Cs in the 12S rRNA gene between nps 956–965 (fig. 1). The MTND1*ADPD3397G mutation has also been found in an independent Alzheimer disease pedigree (Shoffner et al. 1993). Hence, the MTTQ*-ADPD4336C mutation is analogous to the mildest of the LHON mutations, MTND6*LHON14484C and MTCYB*LHON15257A, which together constitute a class of rare mtDNA polymorphisms that predispose individuals to late-onset degenerative diseases (Wallace 1994).

mtDNA Rearrangement Mutations

Over 100 mtDNA rearrangements have been identified in mitochondrial degenerative diseases (Wallace et al. 1995) (fig. 1). These have been associated with three major clinical phenotypes: the ocular myopathies, including the Kearns-Sayre syndrome (KSS) and chronic progressive external ophthalmoplegia (CPEO) (Holt et al. 1988; Moraes et al. 1989; Poulton et al. 1993; Poulton and Holt 1994), Pearson marrow/pancreas syndrome (Rötig et al. 1989, 1990), and adult-onset diabetes mellitus and deafness (Ballinger et al. 1992*b*, 1994). KSS and CPEO are multisystem disorders generally occurring in individuals <20 or >20 years old, respectively, in which the most prominent features are ophthalmoplegia (paralysis of eye muscles), ptosis (droopy eye lids), and mitochondrial myopathy. Additional clinical features include encephalopathy, cardiac conduction defects, renal failure, and diabetes mellitus. Progression of KSS and CPEO have been associated with the progressive increase in the percentage of the rearranged molecules in postmitotic tissues (Mita et al. 1989; Larsson et al. 1990; Shoubridge et al. 1990). Pearson marrow/pancreas syndrome is a frequently lethal childhood disease associated with pancytopenia (loss of all blood cells) due to the accumulation of rearranged mtDNAs in the bone marrow stem cells. Maternally inherited diabetes mellitus and deafness presents with progressive sensory neural hearing loss in the 20s and type II diabetes in the 30s–40s. Severely affected individuals also experience strokelike episodes. In one large three generation family, the inheritance of the diabetes mellitus and deafness phenotype has been associated with the maternal transmission of a trimolecular heteroplasmy involving interrelated normal, duplicated, and deleted molecules. The duplicated molecules contain a 6.1-kb tandem insertion including the control region, rRNA genes, and part of ND1. The deleted molecules share the same breakpoint junction and appear to contain two copies of the 6.1-kb duplicated region but lack the remaining 10.4 kb of the mtDNA (fig. 1). In cultured cells from patients, the duplicated molecules come to predominate, and the deleted molecules are lost, while in skeletal muscle the deleted molecules tend to predominate. This suggests that the normal and duplicated molecules may be maternally transmitted, while the deleted molecules are generated from the duplicated molecules and accumulate in postmitotic tissues, resulting in the onset and progression of the disease (Ballinger et al. 1992*b*, 1994).

Somatic mtDNA Mutations in Progression and Aging

While the age of onset and clinical features of the mitochondrial diseases can be explained by the differing types of mtDNA mutations, an additional factor is required to explain the delayed onset and progression of mitochondrial diseases. This factor has been hypothesized to be the age-related decline of OXPHOS due to the accumulation of somatic mtDNA mutations in postmitotic tissues (Wallace 1992*a*, 1992*b*, and in press; Wallace et al., in press) (fig. 9).

Substantial data have accumulated, indicating that OXPHOS declines with age in a variety of postmitotic tissues in parallel with the accumulation of both mtDNA rearrangements and base-substitution mutations (Wallace in press; Wallace et al., in press). For example, the common 5-kb mtDNA deletion (fig. 1) is essentially absent in normal hearts until age 40 years, after which it progressively accumulates with age (Corral-Debrinski et al. 1991, 1992*b*). Similarly, this deletion is low in all regions of normal brain until age 75 years, after which it accumulates rapidly in the cerebral cortex and basal ganglia, reaching maximum levels of 2%–3% and 11%–12% of total mtDNAs, respectively, by age 80 years. By contrast, this deletion remains at low levels in the cerebellum throughout life (Corral-Debrinski et al. 1992*a*; Soong et al. 1992). Since all mtDNA deletions that have been analyzed have been found to increase concurrently with the 5-kb deletion, current data suggest that a substantial portion of the mtDNAs of postmitotic tissues, such as the basal ganglia, may be mutant in the elderly (Corral-Debrinski et al. 1992*a*, 1992*b*). Hence,

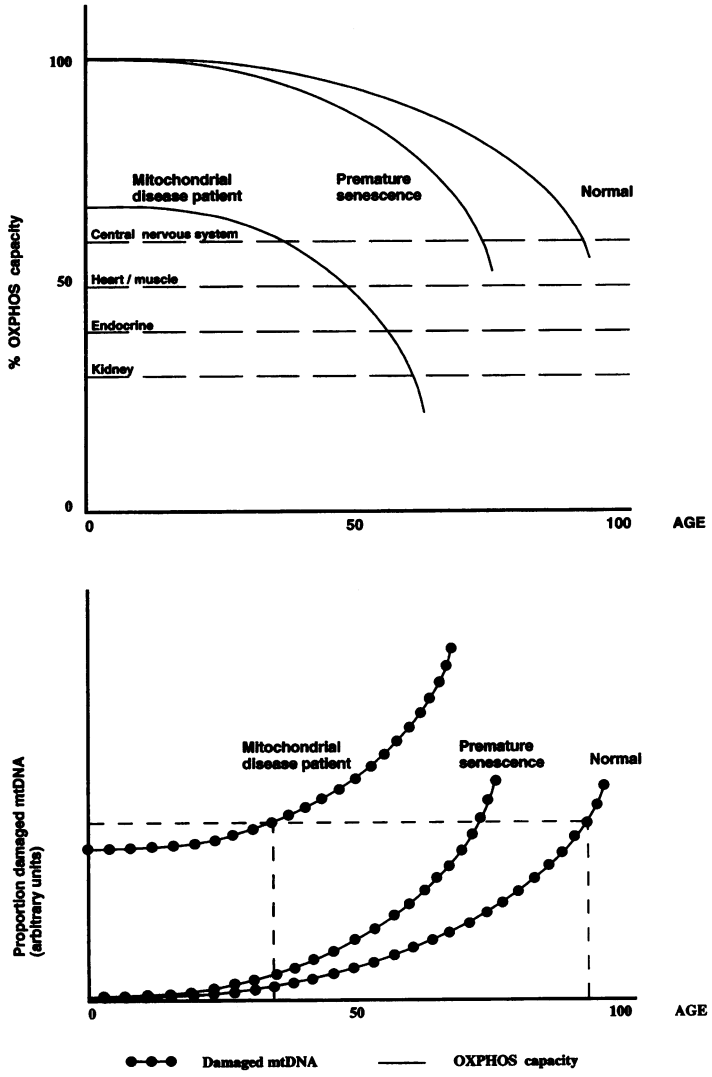

Figure 9 Hypothesis relating the acquisition of mtDNA mutations (inherited and somatic) to the age-related decline of OXPHOS and the progression of OXPHOS diseases and of senescence. The dashed horizontal lines in both panels represent different tissue-specific expression thresholds. The top panel shows the age-related decline of OXPHOS in individuals born with a normal OXPHOS genotype, a mutant OXPHOS gene, and an increased mtDNA somatic mutation rate. The bottom panel shows the relative levels of defective mtDNA with age for each of these individuals.

the age-related accumulation of somatic mtDNA mutations may progressively erode the OXPHOS capacity of postmitotic tissues and thus augment the OXPHOS defect associated with the inherited mtDNA mutations. When the cumulative effect of the inherited plus the somatic mutations reduces the OXPHOS capacity of a tissue to below the bioenergetic expression threshold, symptoms appear and subsequently progress.

If this concept is correct, then individuals born with a normal OXPHOS genotype would start at a high initial bioenergetic capacity and would have to accumulate substantial somatic mtDNA mutations before they would cross expression thresholds. Hence, these individuals would not experience organ degeneration until later in life. By contrast, individuals who inherit a deleterious mtDNA mutation would start with a lower OXPHOS capacity and would require proportionately fewer somatic mtDNA mutations to cross expression thresholds and experience clinical symptoms (Wallace 1992*a*, 1992*b*, and in press; Wallace et al., in press) (fig. 9).

In some diseases, the onset of clinical symptoms may be hastened by an increased rate of somatic mtDNA mutation accumulation. In patients with cardiomyopathy secondary to coronary artery atheroschlerotic plagues, the heart mitochondria are stimulated to generate bursts of oxygen radicals because of cyclic ischemia and reperfusion. This is associated with an 8–2,200-fold increase in somatic mtDNA mutations, presumably

caused by mtDNA mutagenesis from the oxygen radicals (Corral-Debrinski et al. 1991, 1992*b*). Similarly, in late-onset Alzheimer disease, patients analyzed prior to age 75 years have a 15-fold higher cortical somatic deletion level than age-matched controls, while those analyzed after age 75 years have 5-fold less cortical deletion levels. This suggests that many Alzheimer patients have an increased mtDNA somatic mutation rate, and when the neurons accumulate sufficient mutant mtDNAs, they die, resulting in loss of the mutant mtDNAs and associated dementia (Corral-Debrinski et al. 1994). A similar phenomena is seen for Huntington disease patients, who have a 5–11-fold increase in cortical mtDNA deletion levels relative to age-matched controls (Horton et al., in press) (fig. 9).

Conclusion

In summary, mtDNA mutations appear to play a major role in human evolution, degenerative disease, and aging. Neutral and mildly deleterious mtDNA base substitution mutations have been accumulating throughout female history, enhancing human diversity and contributing to our predisposition to postreproductive degenerative diseases. Mildly to severely deleterious base substitution and rearrangement mutations arise repeatly and cause a wide spectrum of degenerative diseases, including blindness, movement disorders, dementias, stroke, heart disease, renal failure, and diabetes. These mutations are frequently heteroplasmic and are removed from the population by selection at a rate corresponding to the severity of the disease. Finally, somatic mtDNA mutations accumulate in postmitotic tissues throughout life, with the rate being accelerated in certain degenerative diseases. These mutations may erode the bioenergetic capacity of tissues causing the onset and progression of maternally-transmitted diseases and contributing to the senescence process in aging. Hence, the novel features of mitochondrial genetics may provide an important explanation for a variety of perplexing phenomena associated with human clinical genetics.

Acknowledgments

This work was supported by National Institutes of Health grants GM46915, NS21328, NS30164, HL45572, AG10130, and DK45215; a Muscular Dystrophy Foundation clinical research grant; and Emory Clinical Research Center grant BR00039.

References

Anderson S, Bankier AT, Barrell BG, de Bruijn MHL, Coulson AR, Drouin J, Eperon IC, et al (1981) Sequence and organization of the human mitochondrial genome. Nature 290:457–465

Bailliet G, Rothhammer F, Carnese FR, Bravi CM, Bianchi NO (1994) Founder mitochondrial haplotypes in Amerindian populations. Am J Hum Genet 55:27–33

Ballinger SW, Schurr TG, Torroni A, Gan YY, Hodge JA, Hassan K, Chen KH, et al (1992*a*) Southeast Asian mitochondrial DNA analysis reveals genetic continuity of ancient mongoloid migrations. Genetics 130:139–152

Ballinger SW, Shoffner JM, Gebhart S, Koontz DA, Wallace DC (1994) Mitochondrial diabetes revisited. Nat Genet 7:458–459

Ballinger SW, Shoffner JM, Hedaya EV, Trounce I, Polak MA, Koontz DA, Wallace DC (1992*b*) Maternally transmitted diabetes and deafness associated with a 10.4-kb mitochondrial DNA deletion. Nat Genet 1:11–15

Blanc H, Adams CW, Wallace DC (1981*a*) Different nucleotide changes in the large rRNA gene of the mitochondrial DNA confer chloramphenicol resistance on two human cell lines. Nucleic Acids Res 9:5785–5795

Blanc H, Chen K-H, D'Amore MA, Wallace DC (1983) Amino acid change associated with the major polymorphic *Hinc*II site of Oriental and Caucasian mitochondrial DNAs. Am J Hum Genet 35:167–176

Blanc H, Wright CT, Bibb MJ, Wallace DC, Clayton DA (1981*b*) Mitochondrial DNA of chloramphenicol-resistant mouse cells contains a single nucleotide change in the region encoding the 3′ end of the large ribosomal RNA. Proc Natl Acad Sci USA 78:3789–3793

Brockington M, Sweeney MG, Hammans SR, Morgan-Hughs JA, Harding AE (1993) A tandem duplication in the D-loop of human mitochondrial DNA is associated with deletions in mitochondrial myopathies. Nat Genet 4:67–71

Brown MD, Torroni A, Reckord CL, Wallace DC. Phylogenetic analysis of Leber's hereditary optic neuropathy mitochondrial DNAs indicates multiple occurrences of the common mutations. Hum Mutat (in press)

Brown MD, Voljavec AS, Lott MT, McDonald I, Wallace DC (1992*a*) Leber's hereditary optic neuropathy: a model for mitochondrial neurodegenerative disease. FASEB J 6:2791–2799

Brown MD, Voljavec AS, Lott MT, Torroni A, Yang C-C, Wallace DC (1992*b*) Mitochondrial DNA complex I and III mutations associated with Leber's hereditary optic neuropathy. Genetics 130:163–173

Brown MD, Wallace DC (1994) Molecular basis of mitochondrial DNA disease. J Bioenerg Biomembr 26:273–289

Brown MD, Yang C-C, Trounce I, Torroni A, Lott MT, Wallace DC (1992*c*) A mitochondrial DNA variant, identified in Leber hereditary optic neuropathy patients, which extends the amino acid sequence of cytochrome c oxidase subunit I. Am J Hum Genet 51:378–385

Brown WM (1980) Polymorphism in mitochondrial DNA of humans as revealed by restriction endonuclease analysis. Proc Natl Acad Sci USA 77:3605–3609

Bunn CL, Wallace DC, Eisenstadt JM (1974) Cytoplasmic inheritance of chlormaphenicol resistance in mouse tissue culture cells. Proc Natl Acad Sci USA 71:1681–1685

——— (1977) Mitotic segregation of cytoplasmic determinants for chloramphenicol resistance in mammalian cells. I. Fusions with mouse cell lines. Somat Cell Genet 3:71–92

Cann RL, Stoneking M, Wilson AC (1987) Mitochondrial DNA and human evolution. Nature 325:31–36

Cann RL, Wilson AC (1983) Length mutations in human mitochondrial DNA. Genetics 104:699–711

Case JT, Wallace DC (1981) Maternal inheritance of mitochondrial DNA polymorphisms in cultured human fibroblasts. Somat Cell Genet 7:103–108

Chen YS, Torroni A, Excoffier L, Santachiara-Benerecetti AS, Wallace DC (1995) Analysis of mtDNA variation in African populations reveals the most ancient of all human continent-specific haplogroups. Am J. Hum Genet 57:133–149

Corral-Debrinski M, Horton T, Lott MT, Shoffner JM, Beal MF, Wallace DC (1992*a*) Mitochondrial DNA deletions in human brain: regional variability and increase with advanced age. Nat Genet 2:324–329

Corral-Debrinski M, Horton T, Lott MT, Shoffner JM, McKee AC, Beal MF, Graham BH, et al (1994) Marked changes in mitochondrial DNA deletion levels in Alzheimer brains. Genomics 23:471–476

Corral-Debrinski M, Shoffner JM, Lott MT, Wallace DC (1992*b*) Association of mitochondrial DNA damage with aging and coronary atherosclerotic heart disease. Mutat Res 275:169–180

Corral-Debrinski M, Stepien G, Shoffner JM, Lott MT, Kanter K, Wallace DC (1991) Hypoxemia is associated with mitochondrial DNA damage and gene induction. JAMA 266:1812–1816

Cortopassi GA, Hutchin TP (1994) Germ-line inheritance of a rare mtDNA variant leads to greatly increased risk for Alzheimer disease. Am J Hum Genet Suppl 55:A149

De Vries DD, Van Engelen BG, Gabreels FJ, Ruitenbeek W, Van Oost BA (1993) A second missense mutation in the mitochondrial ATPase 6 gene in Leigh's syndrome. Ann Neurol 34:410–412

Denaro M, Blanc H, Johnson MJ, Chen KH, Wilmsen E, Cavalli-Sforza LL, Wallace DC (1981) Ethnic variation in *Hpa*I endonuclease cleavage patterns of human mitochondrial DNA. Proc Natl Acad Sci USA 78:5768–5772

DiRienzo A, Wilson AC (1991) Branching pattern in the evolutionary tree for human mitochondrial DNA. Proc Natl Acad Sci USA 88:1597–1601

Giles RE, Blanc H, Cann HM, Wallace DC (1980) Maternal inheritance of human mitochondrial DNA. Proc Natl Acad Sci USA 77:6715–6719

Goto Y, Nonaka I, Horai S (1990) A mutation in the $tRNA^{Leu(UUR)}$ gene associated with the MELAS subgroup of mitochondrial encephalomyopathies. Nature 348:651–653

Greenberg JH, Turner CG II, Zegura SL (1986) The settlement of the Americas: a comparison of the linguistic, dental, and genetic evidence. Curr Anthropol 27:477–497

Hedges SB, Kumar S, Tamura K (1991) Human origins and analysis of mitochondrial DNA sequences. Science 255:737–739

Hertzberg M, Mickleson KNP, Serjeantson SW, Prior JF, Trent RJ (1989) An Asian-specific 9-bp deletion of mitochondrial DNA is frequently found in Polynesians. Am J Hum Genet 44:504–510

Holt IJ, Harding AE, Morgan-Hughes JA (1988) Deletions of muscle mitochondrial DNA in patients with mitochondrial myopathies. Nature 331:717–719

Holt IJ, Harding AE, Petty RK, Morgan-Hughes JA (1990) A new mitochondrial disease associated with mitochondrial DNA heteroplasmy. Am J Hum Genet 46:428–433

Horai S, Hayasaka K (1990) Intraspecific nucleotide sequence differences in the major noncoding region of human mitochondrial DNA. Am J Hum Genet 46:828–842

Horai S, Hayasaka K, Kondo R, Tsugane K, Takahata N (1995) Recent African origin of modern humans revealed by complete sequences of hominoid mitochondrial DNAs. Proc Natl Acad Sci USA 92:532–536

Horai S, Kondo R, Nakagawa-Hattori Y, Hayashi S, Sonoda S, Tajima K (1993) Peopling of the Americas, founded by four major lineages of mitochondrial DNA. Mol Biol Evol 10:23–47

Horai S, Matsunaga E (1986) Mitochondrial DNA polymorphism in Japanese. II. Analysis with restriction enzymes of four or five base pair recognition. Hum Genet 72:105–117

Horton TM, Graham BH, Corral-Debrinski M, Shoffner JM, Kaufman AE, Beal BF, Wallace DC. Marked increase in mitochondrial DNA deletion levels in the cerebral cortex of Huntington's disease patients. Neurology (in press)

Howell N, Bindoff LA, McCullough DA, Kubacka I, Poulton J, Mackey D, Taylor L, et al (1991) Leber hereditary optic neuropathy: identification of the same mitochondrial ND1 mutation in six pedigrees. Am J Hum Genet 49:939–950

Huoponen K, Vilkki J, Aula P, Nikoskelainen EK, Savontaus M-L (1991) A new mtDNA mutation associated with Leber hereditary optic neuroretinopathy. Am J Hum Genet 48:1147–1153

Johns DR, Neufeld MJ (1991) Cytochrome b mutations in Leber hereditary optic neuropathy. Biochem Biophys Res Commun 181:1358–1364

Johns DR, Neufeld MJ, Park RD (1992) An ND-6 mitochondrial DNA mutation associated with Leber hereditary optic neuropathy. Biochem Biophys Res Commun 187:1551–1557

Johnson MJ, Wallace DC, Ferris SD, Rattazzi MC, Cavalli-Sforza LL (1983) Radiation of human mitochondria DNA types analyzed by restriction endonuclease cleavage patterns. J Mol Evol 19:255–271

Jun AS, Brown MD, Wallace DC (1994) A mitochondrial DNA mutation at np 14459 of the ND6 gene associated with maternally inherited Leber's hereditary optic neuropathy and dystonia. Proc Natl Acad Sci USA 91:6206–6210

Kadowaki T, Kadowaki H, Mori Y, Tobe K, Sakuta R, Suzuki Y, Tanabe Y, et al (1994) A subtype of diabetes mellitus associated with a mutation of mitochondrial DNA. N Engl J Med 330:962–968

Kearsey SE, Craig IW (1981) Altered ribosomal RNA genes in mitochondria from mammalian cells with chloramphenicol resistance. Nature 290:607–608

King MP, Attardi G (1988) Injection of mitochondria into human cells leads to a rapid replacement of the endogenous mitochondrial DNA. Cell 52:811–819

Kobayashi Y, Momoi MY, Tominaga K, Momoi T, Nihei K, Yanagisawa M, Kagawa Y, et al (1990) A point mutation in the mitochondrial $tRNA^{Leu(UUR)}$ gene in MELAS (mitochondrial myopathy, encephalopathy, lactic acidosis and stroke-like episodes). Biochem Biophys Res Commun 173:816–822

Larsson NG, Holme E, Kristiansson B, Oldfors A, Tulinius M (1990) Progressive increase of the mutated mitochondrial DNA fraction in Kearns-Sayre syndrome. Pediatr Res 28:131–136

Lucotte G, Guérin P, Hallé L, Loirat F, Hazout S (1989) Y chromosome DNA polymorphisms in two African populations. Am J Hum Genet 45:16–20

Merriwether DA, Clark AG, Ballinger SW, Schurr TG, Soodyall H, Jenkins T, Sherry ST, et al (1991) The structure of human mitochondrial DNA variation. J Mol Evol 33:543–555

Mita S, Schmidt B, Schon EA, DiMauro S, Bonilla E (1989) Detection of "deleted" mitochondrial genomes in cytochrome-c oxidase-deficient muscle fibers of a patient with Kearns-Sayre syndrome. Proc Natl Acad Sci USA 86:9509–9513

Moraes CT, DiMauro S, Zeviani M, Lombes A, Shanske S, Miranda AF, Nakase H, et al (1989) Mitochondrial DNA deletions in progressive external ophthalmoplegia and Kearns-Sayre syndrome. N Engl J Med 320:1293–1299

Neckelmann N, Li K, Wade RP, Shuster R, Wallace DC (1987) cDNA sequence of a human skeletal muscle ADP/ATP translocator: lack of a leader peptide, divergence from a fibroblast translocator cDNA, and coevolution with mitochondrial DNA genes. Proc Natl Acad Sci USA 84:7580–7584

Nei M, Tajima F (1983) Maximum likelihood estimation of the number of nucleotide substitutions from restriction sites data. Genet S Am 105:207–217

Newman NJ, Torroni A, Brown MD, Lott MT, Fernandez MM, Wallace DC, et al (1994) Epidemic neuropathy in Cuba not associated with mitochondrial DNA mutations found in Leber's hereditary optic neuropathy patients. Am J Ophthalmol 118:158–168

Ngo KY, Vergnaud G, Johnsson C, Lucotte G, Weissenbach J (1986) A DNA probe detecting multiple haplotypes of the human Y chromosome. Am J Hum Genet 38:407–418

Novotny EJ, Singh G, Wallace DC, Dorfman LJ, Louis A, Sogg RL, Steinman L (1986) Leber's disease and dystonia: a mitochondrial disease. Neurology 36:1053–1060

Oliver NA, Wallace DC (1982) Assignment of two mitochondrially synthesized polypeptides to human mitochondrial DNA and their use in the study of intracellular mitochondrial interaction. Mol Cell Biol 2:30–41

Ortiz RG, Newman NJ, Shoffner JM, Kaufman AE, Koontz DA, Wallace DC (1993) Variable retinal and neurologic manifestations in patients harboring the mitochondrial DNA 8993 mutation. Arch Opthalmol 111:1525–1530

Persichetti F, Blasi P, Hammer M, Malaspina P, Jodice C, Terrenation L, Novelletto A (1992) Disequilibrium of multiple DNA markers on the human Y chromosome. Ann Hum Genet 56:303–310

Poulton J, Deadman ME, Bindoff L, Morten K, Land J, Brown G (1993) Families of mtDNA re-arrangements can be detected in patients with mtDNA deletions: duplications may be a transient intermediate form. Hum Mol Genet 2:23–30

Poulton J, Holt I (1994) Mitochondrial DNA: does more lead to less? Nat Genet 8:313–315

Rötig A, Colonna M, Bonnefont JP, Blanche S, Fischer A, Saudubray JM, Munnich A (1989) Mitochondrial DNA deletion in Pearson's marrow-pancreas syndrome. Lancet 1:902–903

Rötig A, Cormier V, Blanche S, Bonnefont JP, Ledeist F, Romero N, Schmitz J, et al (1990) Pearson's marrow-pancreas syndrome. A multi-system mitochondrial disorder in infancy. J Clin Invest 86:1601–1608

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Santachiara-Benerecetti AS, Semino O, Passarino G, Torroni A, Brdicka R, Fellous M, Modiano G (1993) The common, near-eastern origin of Ashkenazi and Sephardi Jews supported by Y-chromosome similarity. Ann Hum Genet 57:55–64

Santos M, Barrantes R (1994) D-loop mtDNA deletion as a unique marker of Chibchan Amerindians. Am J Hum Genet 55:413–414

Schurr TG, Ballinger SW, Gan Y-Y, Hodge JA, Merriwether DA, Lawrence DN, Knowler WC, et al (1990) Amerindian mitochondrial DNAs have rare Asian mutations at high frequencies, suggesting they derived from four primary maternal lineages. Am J Hum Genet 46:613–623

Shoffner JM, Bialer MG, Pavlakis SG, Lott MT, Kaufman A, Dixon J, Teichberg S, et al (1995) Mitochondrial encephalomyopathy associated with a single nucleotide deletion in the mitochondrial $tRNA^{Leu(UUR)}$ gene. Neurology 44:286–292

Shoffner JM, Brown MD, Stugard C, Jun AS, Pollok S, Haas RH, Kaufman A, et al (1995) Leber's hereditary optic neuropathy plus dystonia is caused by a mitochondrial DNA point mutation in a complex I subunit. Ann Neurol (in press)

Shoffner JM, Brown MD, Torroni A, Lott MT, Cabell MR, Mirra SS, Beal MF, et al (1993) Mitochondrial DNA variants observed in Alzheimer disease and Parkinson disease patients. Genomics 17:171–184

Shoffner JM, Fernhoff MD, Krawiecki NS, Caplan DB, Holt PJ, Koontz DA, Takei Y, et al (1992) Subacute necrotizing encephalopathy: oxidative phosphorylation defects and the ATPase 6 point mutation. Neurology 42:2168–2174

Shoffner JM, Lott MT, Lezza AM, Seibel P, Ballinger SW, Wallace DC (1990) Myoclonic epilepsy and ragged-red fiber disease (MERRF) is associated with a mitochondrial DNA $tRNA^{Lys}$ mutation. Cell 61:931–937

Shoffner JM, Wallace DC (1994) Oxidative phosphorylation diseases and mitochondrial DNA mutations: Diagnosis and treatment. Ann Rev Nutr 14:535–568

Shoubridge EA, Karpati G, Hastings KEM (1990) Deletion mutants are functionally dominant over wild-type mitochondrial genomes in skeletal muscle fiber segments in mitochondrial disease. Cell 62:43–49

Soong NW, Hinton DR, Cortopassi G, Arnheim N (1992) Mosaicism for a specific somatic mitochondrial DNA mutation in adult human brain. Nat Genet 2:318–323

Spurdle A, Jenkins T (1991) Y chromosome probe p49a detects complex *Pvu*II haplotypes and many new *Taq*I haplotypes in Southern African populations. Am J Hum Genet 50:107–125

Stone AC, Stoneking M (1993) Ancient DNA from a Pre-Columbian Amerindian population. Am J Phys Anthropol 92:463–471

Stoneking M, Jorde LB, Bhatia K, Wilson AC (1990) Geo-

graphic variation in human mitochondrial DNA from Papua New Guinea. Genetics 124:717–733

Swofford DL (1993). Phylogenetic analysis using parsimony (PAUP). Champaign, IL, University of Illinois.

Szathmary EJE (1984) Peopling of northern North America: clues from genetic studies. Acta Anthrop 8:79–109

Tateno Y, Nei M, Tajima F (1982) Accuracy of estimated phylogenetic trees from molecular data. I. Distantly related species. J Mol Evol 18:387–404

Tatuch Y, Christodoulou J, Feigenbaum A, Clarke JTR, Wherret J, Smith C, Rudd N, et al (1992) Heteroplasmic mtDNA mutation (T→G) at 8993 can cause Leigh disease when the percentage of abnormal mtDNA is high. Am J Hum Genet 50:852–858

Templeton AR (1992) Human origins and analysis of mitochondrial DNA sequences. Science 255:737

Torroni A, Brown MD, Lott MT, Newman NJ, Wallace DC, Team CNFI (1995) African, Native American and European mitochondrial DNAs in Cubans from the Pinar del Rio Province and implications for the recent epidemic neuropathy in Cuba. Hum Mutat 5:310–317

Torroni A, Chen Y, Semino O, Santachiara-Beneceretti AS, Scott CR, Lott MT, Winter M, et al (1994*a*) MtDNA and Y-chromosome polymorphisms in four native American populations from southern Mexico. Am J Hum Genet 54:303–318

Torroni A, Lott MT, Cabell MF, Chen Y, Laverge L, Wallace DC (1994*b*) mtDNA and the origin of Caucasians: identification of ancient Caucasian-specific haplogroups, one of which is prone to a recurrent somatic duplication in the D-loop region. Am J Hum Genet 55:760–776

Torroni A, Miller JA, Moore LG, Zamudio S, Zhuang J, Droma R, Wallace DC (1994*c*) Mitochondrial DNA analysis in Tibet. Implications for the origin of the Tibetan population and its adaptation to high altitude. Am J Phys Anthro 93:189–199

Torroni A, Neel JV, Barrantes R, Schurr TG, Wallace DC (1994*d*) A mitochondrial DNA "clock" for the Amerinds and its implication for timing their entry into North America. Proc Natl Acad Sci USA 91:1158–1162

Torroni A, Schurr TG, Cabell MF, Brown MD, Neel JV, Larsen M, Smith DG, et al (1993*a*) Asian affinities and continental radiation of the four founding Native American mtDNAs. Am J Hum Genet 53:563–590

Torroni A, Schurr TG, Yang C-C, Szathmary EJ, Williams RC, Schanfield MS, Troup GA, et al (1992) Native American mitochondrial DNA analysis indicates that the Amerind and the Nadene populations were founded by two independent migrations. Genetics 130:153–162

Torroni A, Semino O, Scozzari R, Sirugo G, Spedini G, Abbas N, Fellous M, et al (1990) Y chromosome DNA polymorphisms in human populations: differences between Caucasoids and Africans detected by 49a and 49f probes. Ann Hum Genet 54:287–296

Torroni A, Sukernik RI, Schurr TG, Starikovskaya YB, Cabell MF, Crawford MH, Comuzzie AG, et al (1993*b*) mtDNA variation of aboriginal Siberians reveals distinct genetic affinities with Native Americans. Am J Hum Genet 53:591–608

Torroni A, Wallace DC (1995) Mitochondrial DNA haplogroups in Native Americans. Am J Hum Genet 56:1234–1236

Trounce I, Neill S, Wallace DC (1994) Cytoplasmic transfer of the mtDNA nt 8993 T→G (ATP6) point mutation associated with Leigh syndrome into mtDNA-less cells demonstrates cosegregation with a decrease in state III respiration and ADP/O ratio. Proc Natl Acad Sci USA 91:8334–8338

Turner CG II (1983) Dental evidence for the peopling of the Americas. In: Shutler R Jr. (ed) Early man in the new world. Sage, Beverly Hills, CA, pp 147–157

——— (1987) Late Pleistocene and Holocene population history of East Asia based on dental variation. Am J Phys Anthropol 73:305–321

van den Ouweland JM, Lemkes HH, Trembath RC, Ross R, Velho G, Cohen D, Froguel P, et al (1994) Maternally inherited diabetes and deafness is a distinct subtype of diabetes and associates with a single point mutation in the mitochondrial $tRNA^{Leu(UUR)}$ gene. Diabetes 43:746–51

van den Ouweland JM, Lemkes HHP, Ruitenbeek W, Sandkjujl LA, deVijlder MF, Struyvenberg PAA, van de Kamp JJP, et al (1992) Mutation in mitochondrial $tRNA^{Leu(UUR)}$ gene in a large pedigree with maternally transmitted type II diabetes mellitus and deafness. Nat Genet 1:368–371

Vigilant L, Stoneking M, Harpending H, Hawkes K, Wilson AC (1991) African populations and the evolution of human mitochondrial DNA. Science 253:1503–1507

Wallace DC (1981) Assignment of the chloramphenicol resistance gene to mitochondrial deoxyribonucelic acid and analysis of its expression in cultured human cells. Mol Cell Biol 1:697–710

——— (1982*a*) Cytoplasmic inheritance of chloramphenicol resistance in mammalian cells. In: Shay JW (ed) Techniques in somatic cell genetics. Plenum Press, New York, pp 159–187

——— (1982*b*) Structure and evolution of organelle genomes. Microbiol Rev 46:208–240

——— (1986) Mitotic segregation of mitochondrial DNAs in human cell hybrids and expression of chloramphenicol resistance. Somat Cell Mol Genet 12:41–49

——— (1992*a*) Diseases of the mitochondrial DNA. Annu Rev Biochem 61:1175–1212

——— (1992*b*) Mitochondrial genetics: a paradigm for aging and degenerative diseases? Science 256:628–632

——— (1994) Mitochondrial DNA sequence variation in human evolution and disease. Proc Natl Acad Sci USA 91:8739–8746

———. Mitochondrial DNA Mutations in Human Disease and Aging. In: Esser K, Martin GM (eds) Molecular aspects of aging. John Wiley & Sons, New York (in press)

Wallace DC, Bunn CL, Eisenstadt JM (1975) Cytoplasmic transfer of chloramphenicol resistance in human tissue culture cells. J Cell Biol 67:174–188

——— (1977) Mitotic segregation of cytoplasmic inherited determinants for chloramphenicol resistance in mammalian cells. II. Fusions with human cell lines. Somat Cell Genet 3:93–119

Wallace DC, Garrison K, Knowler WC (1985) Dramatic founder effects in Amerindian mitochondrial DNAs. Am J Phys Anthropol 68:149–155

Wallace DC, Lott MT (1992) Maternally inherited diseases. In: DiMauro Wallace DC (ed) Mitochondrial DNA in human pathology. Raven Press, New York, pp 63–83

Wallace DC, Lott MT, Brown MD, Huoponen K, Torroni A (1995) Report of the committee on human mitochondrial DNA. In: Cuticchia AJ (ed) Human gene mapping 1995: a compendium. Johns Hopkins University Press, Baltimore, pp 910–954 (also available at http://infinity.gen.amory.edu/mitomap.html)

Wallace DC, Lott MT, Shoffner JM, Ballinger S (1994) Mitochondrial DNA mutations in epilepsy and neurological disease. Epilepsia Suppl 35:S43–S50

Wallace DC, Richter C, Bohr VA, Cortopassi G, Kadenbach B, Linn S, Linnane AW, et al. Group report: the role of bioenergetics and mitochondrial DNA mutations in aging and age-related diseases. In: Esser K, Martin GM (eds) Molecular aspects of aging. John Wiley & Sons, New York (in press)

Wallace DC, Singh G, Lott MT, Hodge JA, Schurr TG, Lezza AM, Elsas LJ, et al (1988*a*) Mitochondrial DNA mutation associated with Leber's hereditary optic neuropathy. Science 242:1427–1430

Wallace DC, Torroni A (1992) American Indian prehistory as written in the mitochondrial DNA: a review. Hum Biol 64:403–416

Wallace DC, Ye JH, Neckelmann SN, Singh G, Webster KA, Greenberg BD (1987) Sequence analysis of cDNAs for the human and bovine ATP synthase β-subunit: mitochondrial DNA genes sustain seventeen times more mutations. Curr Genet 12:81–90

Wallace DC, Zheng X, Lott MT, Shoffner JM, Hodge JA, Kelley RI, Epstein CM, et al (1988*b*) Familial mitochondrial encephalomyopathy (MERRF): genetic, pathophysiological, and biochemical characterization of a mitochondrial DNA disease. Cell 55:601–610

Ward RH, Frazier BL, Dew-Jager K, Paabo S (1991) Extensive mitochondrial diversity within a single Amerindian tribe. Proc Natl Acad Sci USA 88:8720–8724

Ward RH, Redd A, Valencia D, Frazier B, Paabo S (1993) Genetic and linguistic differentiation in the Americas. Proc Natl Acad Sci USA 90:10663–10667

Wrischnik LA, Higuchi RG, Stoneking M, Erlich HA, Arnheim N, Wilson AC (1987) Length mutations in human mitochondrial DNA: direct sequencing of enzymatically amplified DNA. Nucleic Acids 15:529–542